A joint normal-binary(probit) model

Margaux Delporte[1] **Steffen Fieuws**[1] **Geert Molenberghs**[2,1] **Geert Verbeke**[1,2]

Simeon Situma Wanyama[3] **Elpis Hatziagorou**[4] **Christiane De Boeck**[5]

[1] *I-BioStat, KU Leuven, B-3000 Leuven, Belgium*

[2] I-BioStat, Universiteit Hasselt, B-3590 Diepenbeek, Belgium

[3] *Department of Epidemiology and public health, Sciensano, BE-1050 Brussels, Belgium*

[4] Paediatric Pulmonology and CF Unit, Hippokration Hospital of Thessaloniki, *Aristotle University of Thessaloniki, Thessaloniki, Greece*

[5] University Hospital Gasthuisberg, University of Leuven, Leuven, Belgium

Supplementary Materials

A Computation of the marginal joint longitudinal probit-normal model

We will sketch the derivation of (8). In order to do so, the random effects $\boldsymbol{b}_i$ are integrated out of the joint density of $\boldsymbol{y}_{1i}$, $\boldsymbol{y}_{2i}$ and $\boldsymbol{b}_i$.
Because the elements of $\boldsymbol{y}_{1i}$ and $\boldsymbol{y}_{2i}$ are independently normally distributed, conditional on the random effects, we have for the marginal probability:

$$
\begin{aligned}
& f(\boldsymbol{y}_{1i}, \boldsymbol{y}_{2i} = 1) \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \int_{t=-\infty}^{t=X_{2i}\beta+Z_{2i}\boldsymbol{b}_i} \frac{1}{(2\pi)^{(q+n_i+p_i)/2}|D|^{1/2}|\Sigma_i|^{1/2}} \text{exp}\Big\{-\frac{1}{2}\big[\boldsymbol{b}_i' D^{-1}\boldsymbol{b}_i\big]\Big\} \\
&\quad \text{exp}\Big\{-\frac{1}{2}\big[(\boldsymbol{y}_{1i} - X_{1i}\beta - Z_{1i}\boldsymbol{b}_i)'\Sigma_i^{-1}(\boldsymbol{y}_{1i} - X_{1i}\beta - Z_{1i}\boldsymbol{b}_i) + \boldsymbol{t}'\boldsymbol{t}\big]\Big\} d\boldsymbol{b}_i \text{d}\boldsymbol{t} \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=X_{2i}\beta} \frac{1}{(2\pi)^{(q+n_i+p_i)/2}|D|^{1/2}|\Sigma_i|^{1/2}} \text{exp}\Big\{-\frac{1}{2}\big[\boldsymbol{b}_i' D^{-1}\boldsymbol{b}_i\big]\Big\} \\
&\quad \text{exp}\Big\{-\frac{1}{2}\big[(\gamma_i - Z_{1i}\boldsymbol{b}_i)'\Sigma_i^{-1}(\gamma_i - Z_{1i}\boldsymbol{b}_i) + (\boldsymbol{s} + Z_{2i}\boldsymbol{b}_i)'(\boldsymbol{s} + Z_{2i}\boldsymbol{b}_i)\big]\Big\} d\boldsymbol{b}_i \text{d}\boldsymbol{s} \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=X_{2i}\beta} \frac{1}{(2\pi)^{(q+n_i+p_i)/2}|D|^{1/2}|\Sigma_i|^{1/2}} \text{exp}\Big\{-\frac{1}{2}\big[\boldsymbol{b}_i - u_i)'K_i^{-1}(\boldsymbol{b}_i - u_i) + V_i\big]\Big\} d\boldsymbol{b}_i \text{d}\boldsymbol{s}.
\end{aligned}
$$

Here, we have rewritten the variable $\boldsymbol{t} = \boldsymbol{s} + Z_{2i}\boldsymbol{b}_i$ and $\gamma_i = \boldsymbol{y}_{1i} - X_{1i}\beta$. Further,

$$\boldsymbol{b}_i' D^{-1}\boldsymbol{b}_i + (\gamma_i + Z_{1i}\boldsymbol{b}_i)'\Sigma_i^{-1}(\gamma_i + Z_{1i}\boldsymbol{b}_i) + (\boldsymbol{s} + Z_{2i}\boldsymbol{b}_i)'(\boldsymbol{s} + Z_{2i}\boldsymbol{b}_i) = (\boldsymbol{b}_i - u_i)'K_i^{-1}(\boldsymbol{b}_i - u_i) + \nu_i,$$

where

$$
\begin{aligned}
u_i &= -K_i l_i, \\
K_i^{-1} &= D^{-1} + Z_{1i}' \Sigma_i^{-1} Z_{1i} + Z_{2i}' Z_{2i}, \\
l_i' &= -\gamma_i' \Sigma_i^{-1} Z_{1i} + \boldsymbol{s}' Z_{2i}, \\
\nu_i &= \gamma_i' \Sigma_i^{-1} \gamma_i - (Z_{2i}' \boldsymbol{s} - Z_{1i}' \Sigma_i^{-1} \gamma_i)' K_i (Z_{2i}' \boldsymbol{s} - Z_{1i}' \Sigma_i^{-1} \gamma_i) + \boldsymbol{s}' \boldsymbol{s}.
\end{aligned}
$$

Next, integration over the random effects produces

$$
\begin{aligned}
& f(\boldsymbol{y}_{1i}, \boldsymbol{y}_{2i} = \mathbf{1}) \\
&= \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=X_{2i}\beta} \frac{|K_i|^{1/2}}{|D|^{1/2} |\Sigma_i|^{1/2} (2\pi)^{(p_i+n_i)/2}} \mathsf{exp}\{-\frac{1}{2}\nu_i\} \mathsf{d}\boldsymbol{s} \\
&= \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=X_{2i}\beta} \frac{|K_i|^{1/2} |B_i|^{1/2}}{|D|^{1/2} (2\pi)^{(p_i+n_i)/2} |\Sigma_i|^{1/2} |B_i|^{1/2}} \mathsf{exp}\Big\{ -\frac{1}{2} \Big[(s-\alpha_i)' B_i^{-1} (s-\alpha_i) + c_i \Big] \Big\} \mathsf{d}\boldsymbol{s},
\end{aligned}
$$

where $\nu_i = (s-\alpha_i)' B_i^{-1} (s-\alpha_i) + c_i$, with

$$
\begin{aligned}
\alpha_i &= -B_i Z_{2i} K_i Z_{1i}' \Sigma_i^{-1} \gamma_i, \\
B_i^{-1} &= I - Z_{2i} K_i Z_{2i}', \\
c_i &= -\alpha_i' B_i^{-1} \alpha_i + \gamma_i' (\Sigma_i^{-1} - \Sigma_i^{-1\prime} Z_{1i} K_i Z_{1i}' \Sigma_i^{-1}) \gamma_i.
\end{aligned}
$$

Further, write $s - \alpha_i = u$, which results in

$$
f(\boldsymbol{y}_{1i}, \boldsymbol{y}_{2i} = \mathbf{1}) = \frac{|K_i|^{1/2} |B_i|^{1/2}}{|D|^{1/2} (2\pi)^{n_i/2} |\Sigma_i|^{1/2}} \mathsf{exp}(\frac{-c_i}{2}) \Phi(X_{2i}\beta - \alpha_i, B_i).
$$

Now, consider: $\begin{bmatrix} K_i^{-1} & Z_{21i}' \\ Z_{21i} & I \end{bmatrix}$

and

Then

$$
\begin{aligned}
|K_i|^{-1} \cdot |I - Z_{2i} K_i Z_{2i}'| &= |I| \cdot |K_i^{-1} - Z_{2i}' Z_{2i}| \\
|K_i|^{-1} |B_i|^{-1} &= |D^{-1} + Z_{1i}' \Sigma_i^{-1} Z_{1i}| \\
|K_i|^{1/2} |B_i|^{1/2} &= |D^{-1} + Z_{1i}' \Sigma_i^{-1} Z_{1i}|^{-1/2}.
\end{aligned}
$$

Applying this result produces

$$
f(\boldsymbol{y}_{1i}, \boldsymbol{y}_{2i} = \mathbf{1}) = \frac{|D^{-1} + Z_{1i}' \Sigma_i^{-1} Z_{1i}|^{-1/2}}{|D|^{1/2} (2\pi)^{n_i/2} |\Sigma_i|^{1/2}} \mathsf{exp}(\frac{-c_i}{2}) \Phi_{p_i}(X_{2i}\beta - \alpha_i, B_i).
$$

Now write

$$
\begin{aligned}
c_i &= \gamma_i' (\Sigma_i^{-1} - \Sigma_i^{-1\prime} Z_{1i} K_i (K_i^{-1} + Z_{2i}' B_i Z_{2i}) K_i Z_{1i}' \Sigma_i^{-1}) \gamma_i, \\
c_i &= \gamma_i' W_i^{-1} \gamma, \\
W_i^{-1} &= \Sigma_i^{-1} - \Sigma_i^{-1} Z_{1i} K_i [K_i^{-1} + Z_{2i}' B_i Z_{2ij}] K_i Z_{1i}' \Sigma_i^{-1}.
\end{aligned}
$$

Further, consider $\begin{bmatrix} -D^{-1} & Z'_{1i} \\ Z_{1i} & I \end{bmatrix}$

Hence,

$$|D^{-1}| \cdot |\Sigma_i + Z_{1i} D Z'_{1i}| = |\Sigma_i| \cdot |D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i}|.$$

Hence,

$$\frac{1}{|D|^{1/2} |\Sigma_i|^{1/2} |D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i}|^{1/2}} = \frac{1}{|\Sigma_i + Z_{1i} D Z'_{1i}|^{1/2}}.$$

Applying this result results in

$$f(\boldsymbol{y}_{1i}, \boldsymbol{y}_{2i} = \mathbf{1}) = \frac{|W_i|^{1/2}}{|V_i|^{1/2}} \phi(X_{1i}\beta; W_i) \Phi(X_{2i}\beta - \alpha_i; B_i).$$

Now, consider

$$\begin{aligned} D^{-1} + Z'_{2i} Z_{2i} + Z'_{1i} \Sigma_i^{-1} Z_{1i} - Z'_{2i} Z_{2i} &= D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i} \\ D^{-1} + Z'_{2i} Z_{2i} + Z'_{1i} \Sigma_i^{-1} Z_{1i} - Z'_{2i} (I - Z_{2i} K_i Z'_{2i} + Z_{2i} K_i Z'_{2i})^{-1} Z_{2i} &= D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i}. \end{aligned}$$

Inserting $B_i^{-1} = I - Z_{2i} K_i Z'_{2i}$ and $K^{-1} = D^{-1} + Z'_{2i} Z_{2i} + Z'_{1i} \Sigma_i^{-1} Z_{1i}$ results in

$$K^{-1} - Z'_{2i} (B_i^{-1} + Z_{2i} K_i Z'_{2i})^{-1} Z_{2i} = D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i}.$$

Next, taking the inverse of both sides results in

$$\begin{aligned} K_i + K_i Z'_{2i} B_i Z_{2i} K_i &= (D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i})^{-1} \\ K_i [K_i^{-1} + Z'_{2i} B_i Z_{2i}] K_i &= (D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i})^{-1} \\ \Sigma_i^{-1} - \Sigma_i^{-1} Z_{1i} K_i [K_i^{-1} + Z'_{2i} B_i Z_{2i}] K_i Z'_{1i} \Sigma_i^{-1} &= \Sigma_i^{-1} - \Sigma_i^{-1} Z_{1i} (D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i})^{-1} Z'_{1i} \Sigma_i^{-1}. \end{aligned}$$

Since $V^{-1} = \Sigma_i^{-1} - \Sigma_i^{-1} Z_{1i} (D^{-1} + Z'_{1i} \Sigma_i^{-1} Z_{1i})^{-1} Z'_{1i} \Sigma_i^{-1}$, this results in

$$V^{-1} = W^{-1}.$$

Applying this result produces

$$f(\boldsymbol{y}_{1i}, \boldsymbol{y}_{2i} = \mathbf{1}) = \phi(X_{1i}\beta; V_i) \Phi(X_{2i}\beta - \alpha_i; B_i).$$

Note that (8) defines the joint probability of the continuous response with a vector of successes of the binary response $\boldsymbol{y}_{2i} = (1, 1, .., 1)'$. One can derive the joint probability with an arbitrary vector of successes and failures by applying appropriate contrasts. Precisely,

$$f(\boldsymbol{y}_{1i}, \boldsymbol{y}_{2i} = \boldsymbol{m}_i = (m_{i1}, \ldots, m_{ip_i})']) = \phi(X_{1i}\beta; V_i) \Phi(\boldsymbol{s} \diamond (X_{2i}\beta - \alpha_i); C \diamond B_i), \qquad (16)$$

Where $\diamond$ indicates pairwise multiplication. Further,

$$s_{ik} = \begin{cases} 1 & \text{if } m_{ik} \text{ is 1,} \\ -1 & \text{otherwise,} \end{cases}$$

$$C_{ikl} = \begin{cases} 1 & \text{if } m_{ik} = m_{il}, \\ -1 & \text{otherwise.} \end{cases}$$

B Conditional distribution of the the continuous response given the binary response

B.1 Expected value

Let us derive (10), the conditional expected value of the $\widetilde{n}_i$-dimensional continuous subvector $\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}$ given the $\widetilde{p}_i$-dimensional binary subvector $\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}}$. This is equal to the integral over $\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}$ multiplied by the conditional distribution, which is defined as the quotient of (8) and (4).

$$
\begin{aligned}
&E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=1]\\
&= \int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=\infty} \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} \frac{\phi(X_{1i}\beta;V_i)\Phi(X_{2i}\beta-\alpha_i;B_i)}{\Phi(\widetilde{X}_{2i}\beta,L^{-1})} d\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\\
&= c\int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=\infty}\int_{\boldsymbol{t}=-\infty}^{\boldsymbol{t}=\widetilde{X}_{2i}\beta-H_i\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}+H_i\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(\widetilde{n}_i+\widetilde{p}_i)}{2}}}\frac{1}{\sqrt{|V_i||B_i|}}\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\\
&\quad\times \exp\left\{-\frac{1}{2}\left((\widetilde{Y}-\widetilde{X}_{1i}\beta)'V_i^{-1}(\widetilde{Y}-\widetilde{X}_{1i}\beta)+(t'B_i^{-1}t)\right)\right\}d\boldsymbol{t}\; \mathrm{d}\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\\
&= c\int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=\infty}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H_i\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(\widetilde{n}_i+\widetilde{p}_i)}{2}}}\frac{1}{\sqrt{|V_i||B_i|}}\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\\
&\quad\times \exp\left\{-\frac{1}{2}\left((\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}-\widetilde{X}_{1i}\beta)'V_i^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}-\widetilde{X}_{1i}\beta)+(s-H_i\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}})'B_i^{-1}(s-H_i\boldsymbol{y}_{1i})\right)\right\}dsd\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\\
&= c\int_{-\infty}^{\infty}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H_i\widetilde{X}_{1i}\beta}\frac{1}{(2\pi)^{\frac{(\widetilde{n}_i+\widetilde{p}_i)}{2}}}\frac{1}{\sqrt{|V_i||B_i|}}\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\exp\left\{-\frac{1}{2}\left((\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}-U_i)'E_i^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}-U_i)+O_i\right)\right\}dsd\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}},
\end{aligned}
$$

where

$$
\begin{aligned}
H_i &= -B_i\widetilde{Z}_{2i}'K_i\widetilde{Z}_{1i}\Sigma_i^{-1}\\
c &= \frac{1}{\Phi(\widetilde{X}_{2i}\beta,L^{-1})}.
\end{aligned}
$$

In addition, we have substituted $s=t-H_i\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}$, and further

$$
(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}-\widetilde{X}_{1i}\beta)'V_i^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}-\widetilde{X}_{1i}\beta)+(s-H_i\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}})'B_i^{-1}(s-H_i\boldsymbol{y}_{1i})=(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}-U_i)'E_i^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}-U_i)+O_i),
$$

where

$$
\begin{aligned}
E_i^{-1} &= H_i'B_i^{-1}H_i+V_i^{-1}\\
l_i' &= -\boldsymbol{s}'B_i^{-1}H_i-\widetilde{X}_{1i}\beta_1'V_i^{-1}\\
O_i &= \boldsymbol{s}'B_i^{-1}\boldsymbol{s}+\widetilde{X}_{1i}\beta_1'V_i^{-1}\widetilde{X}_{1i}\beta_1-(-H_i'B_i^{-1}\boldsymbol{s}-V_i^{-1}\widetilde{X}_{1i}\beta_1)'E(-H_i'B_i^{-1}\boldsymbol{s}-V_i^{-1}\widetilde{X}_{1i}\beta_1)\\
U_i &= -E_il_i.
\end{aligned}
$$

Integration over $\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}$ produces

$$
\begin{aligned}
&E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=1]\\
&= c\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H_i\widetilde{X}_{1i}\beta}\frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\frac{|E_i|^{1/2}}{\sqrt{|V_i||B_i|}}E_i(H_i'B_i^{-1}s+V_i^{-1}\widetilde{X}_{1i}\beta)\mathrm{exp}\Big\{-\frac{1}{2}O_i\Big)\Big\}d\boldsymbol{s}\\
&= c\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H_i\widetilde{X}_{1i}\beta}\frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\frac{|E_i|^{1/2}}{\sqrt{|V_i||B_i|}}E_i(H_i'B_i^{-1}s+V_i^{-1}\widetilde{X}_{1i}\beta)\mathrm{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s}-F_i)'T_i^{-1}(\boldsymbol{s}-F_i)+Gi\Big)\Big\}d\boldsymbol{s}\\
&= \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H_i\widetilde{X}_{1i}\beta}\frac{c}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\frac{|E_i|^{1/2}}{\sqrt{|V_i||B_i|}}E_iH_i'B_i^{-1}s\mathrm{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s}-F_i)'T_i^{-1}(\boldsymbol{s}-F_i)+Gi\Big)\Big\}d\boldsymbol{s}\\
&\quad +c\ E_iV_i^{-1}\widetilde{X}_{1i}\beta\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H_i\widetilde{X}_{1i}\beta}\frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\mathrm{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s}-F_i)'T_i^{-1}(\boldsymbol{s}-F_i)+Gi\Big)\Big\}d\boldsymbol{s},
\end{aligned}
$$

where we have subsituted $O_i=(\boldsymbol{s}-F_i)'T_i^{-1}(\boldsymbol{s}-F_i)+G_i$,

with

$$
\begin{aligned}
T_i^{-1} &= B_i^{-1}-(H_i'B_i^{-1})'E_i(H_i'B_i^{-1})\\
F_i &= T_i\cdot(H_i'B_i^{-1})'E_i(V_i^{-1}\widetilde{X}_{1i}\beta)\\
G_i &= -F_i'T_i^{-1}F_i+(\widetilde{X}_{1i}\boldsymbol{\beta})'V_i^{-1}(\widetilde{X}_{1i}\boldsymbol{\beta})-(V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta})'E_i(V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}).
\end{aligned}
$$

In order to solve the first integral, we use the formula of the expectation of the truncated normal distribution, described by Manjunath and Wilhelm (2012) in their Equation 17.

$$
\begin{aligned}
&= c\ \mathrm{exp}\{-\frac{1}{2}G_i\}\sqrt{\frac{|E_i||T_i|}{|V_i||B_i|}}E_iH_i'B_i^{-1}T_i\Phi(o,F_i,T_i)\big[-F_1(o_1)\quad -F_2(o_2)\quad \dots\quad -F_p(o_p)\big] \quad (17)\\
&\quad +c\ \mathrm{exp}\{-\frac{1}{2}G_i\}\sqrt{\frac{|E_i||T_i|}{|V_i||B_i|}}\Phi(\widetilde{X}_{2i}\beta+H_i\widetilde{X}_{1i}\beta,F_i,T_i)E_iH_i'B_i^{-1}F_i\\
&\quad +c\ \mathrm{exp}\{-\frac{1}{2}G_i\}\sqrt{\frac{|E_i||T_i|}{|V_i||B_i|}}\Phi(\widetilde{X}_{2i}\beta+H_i\widetilde{X}_{1i}\beta,F_i,T_i)E_iV_i^{-1}\widetilde{X}_{1i}\beta,
\end{aligned}
$$

where o,$F_i(x_i)$ and $\varphi(x_i)$ are defined in (18).

It is important to note that (10) is conditional on a vector of successes (1,1,...,1)'. In analogy with this derivation, we can derive the expected value given an arbitrary vector of successes and failures by using appropriate contrasts. The conditional expected value given an arbitrary vector is defined as

$$
\begin{aligned}
E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{2i}\quad &= \boldsymbol{m}_i=(m_{i1},\dots,m_{ip_i})']]=\frac{\mathrm{exp}\{-\frac{1}{2}G_i\}}{\Phi(\widetilde{X}_{2i}\beta^*,(L)^{-1})}\sqrt{\frac{|E_i||T_i|}{|V_i||B_i|}}\Phi(\widetilde{X}_{2i}\beta^*+H_i\widetilde{X}_{1i}\beta,F_i,T_i) \quad (18)\\
&\times\Big(E_i(V_i^{-1}\widetilde{X}_{1i}\beta+H_i'B_i^{-1}F_i)-E_iH_i'B_i^{-1}T_i\big[-F_1(o_1)\quad -F_2(o_2)\quad \dots\quad -F_p(o_p)\big]\Big),
\end{aligned}
$$

with

$$\widetilde{X}_{2i}\beta^* = s \diamond \widetilde{X}_{2i}\beta$$
$$\widetilde{Z}_{2i}^* = J \diamond Z_{2i},$$

where $\diamond$ signals pairwise multiplication and s is defined in (16). Further

$$J_{ikl} = \begin{cases} 1 & \text{if } m_{ik} \text{ is } 1, \\ -1 & \text{otherwise.} \end{cases}$$

B.2 Prediction interval

The prediction interval of the expected values of a (sub)vector of the continuous response given a (sub)vector of the binary response is composed by the variability of the observations (second central moment) and the standard errors of the transformed parameters. We will first derive the second central moment and then derive the standard errors via the delta method.

The uncertainty of a new observation is defined as the second central moment

$$
\begin{aligned}
&E\Big[\big(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\big)\big(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\big)'\Big] \\
&= \int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=+\infty} (\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})](\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}])' \cdot f(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}) d\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} \\
&= \int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=+\infty} \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}{}' f(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}) - E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})' f(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}) \\
&\quad -(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}{}' E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}) + E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}{}') \cdot f(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}) d\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} \\
&= \int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=+\infty} \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}{}' f(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}) d\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})' \\
&= c\int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=+\infty} \int_{\boldsymbol{t}=-\infty}^{\boldsymbol{t}=\widetilde{X}_{2i}\beta - H\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(\tilde{n}_i+\tilde{p}_i)}{2}}} \frac{1}{\sqrt{|V_i||B_i|}} \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}'_{\boldsymbol{1i}} \\
&\quad \times \exp\bigg\{ -\frac{1}{2}\Big((\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\beta)' V_i^{-1} (\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\beta) + (\boldsymbol{t}' B_i^{-1} \boldsymbol{t}) \Big) \bigg\} d\boldsymbol{t} d\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})' \\
&= c\int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=+\infty} \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(\tilde{n}_i+\tilde{p}_i)}{2}}} \frac{1}{\sqrt{|V_i||B_i|}} \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}'_{\boldsymbol{1i}} \\
&\quad \times \exp\bigg\{ -\frac{1}{2}\Big((\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\beta)' V_i^{-1} (\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\beta) + ((\boldsymbol{s} - H\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}})' B_i^{-1} (\boldsymbol{s} - H\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}})) \Big) \bigg\} d\boldsymbol{s} d\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} \\
&\quad - E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})' \\
&= c\int_{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}=+\infty} \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(\tilde{n}_i+\tilde{p}_i)}{2}}} \frac{1}{\sqrt{|V_i||B_i|}} \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}'_{\boldsymbol{1i}} \\
&\quad \times \exp\bigg\{ -\frac{1}{2}\Big((\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - u_i)' E_i^{-1} (\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - u_i) + O_i \Big) \bigg\} d\boldsymbol{s} d\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})'
\end{aligned}
$$

where we have substituted $s = t - H\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}$, and further

$$(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\beta)'V_i^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\beta) + (s - H\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}})'B_i^{-1}(s - H\boldsymbol{y}_{1i}) = (\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - U_i)'E_i^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} - U_i) + O_i,$$

where

$$\begin{aligned}
E_i^{-1} &= H'B_i^{-1}H + V_i^{-1} \\
l_i' &= -\boldsymbol{s}'B_i^{-1}H - \widetilde{X}_{1i}\beta_1'V_i^{-1} \\
O_i &= \boldsymbol{s}'B_i^{-1}\boldsymbol{s} + \widetilde{X}_{1i}\beta_1'V_i^{-1}\widetilde{X}_{1i}\beta_1 - (-H'B_i^{-1}\boldsymbol{s} - V_i^{-1}\widetilde{X}_{1i}\beta_1)'E(-H'B_i^{-1}\boldsymbol{s} - V_i^{-1}\widetilde{X}_{1i}\beta_1) \\
U_i &= -E_i l_i.
\end{aligned}$$

Now we can integrate over $\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}$, applying the formula of the second moment of a multivariate gaussian distribution:

$$\begin{aligned}
&E\Big[\big(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\big)^2\Big] \\
&= c\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\frac{|E_i|^{1/2}}{\sqrt{|V_i||B_i|}}u_iu_i'\text{exp}\Big\{-\frac{1}{2}O_i\Big\}d\boldsymbol{s} \\
&\quad +E_ic\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\frac{|E_i|^{1/2}}{\sqrt{|V_i||B_i|}}\text{exp}\Big\{-\frac{1}{2}O_i\Big\}d\boldsymbol{s} - E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})' \\
&= c\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\frac{|E_i|^{1/2}}{\sqrt{|V_i||B_i|}}u_iu_i'\text{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s}-F_i)'T_i^{-1}(\boldsymbol{s}-F_i)+G_i\Big)\Big\}d\boldsymbol{s} \\
&\quad +E_ic\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\frac{|E_i|^{1/2}}{\sqrt{|V_i||B_i|}}\text{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s}-F_i)'T_i^{-1}(\boldsymbol{s}-F_i)+G_i\Big)\Big\}d\boldsymbol{s} \\
&\quad -E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})',
\end{aligned}$$

where we have rewritten $O_i = (s - F_i)'T_i^{-1}(s - F_i) + G_i$, with

$$\begin{aligned}
T_i^{-1} &= B_i^{-1} - (H'B_i^{-1})'E_i(H'B_i^{-1}) \\
F_i &= T_i \cdot (H'B_i^{-1})E_i(V_i^{-1}\widetilde{X}_{1i}\beta) \\
G_i &= -F_i'T_i^{-1}F_i + (\widetilde{X}_{1i}\boldsymbol{\beta})'V_i^{-1}(\widetilde{X}_{1i}\boldsymbol{\beta}) - (V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta})'E_i(V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}).
\end{aligned}$$

Integration over $\boldsymbol{s}$ produces

$$E\Big[\big(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}-E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\boldsymbol{1}]\big)\big(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}-E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\boldsymbol{1}]\big)'\Big]$$

$$\begin{aligned}
&= c\frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}}\frac{|E_i|^{1/2}}{\sqrt{|V_i||B_i|}}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta}\Big(E_iH'B^{-1}\boldsymbol{s}\boldsymbol{s}'B^{-1}HE_i+E_iH'B^{-1}\boldsymbol{s}(\widetilde{X}_{1i}\beta)'V_i^{-1}E_i \qquad (19)\\
&\quad +EV_i^{-1}\widetilde{X}_{1i}\beta\boldsymbol{s}'B^{-1}HE_i+EV_i^{-1}\widetilde{X}_{1i}\beta(\widetilde{X}_{1i}\beta)'V_i^{-1}E_i)\Big)\text{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s}-F_i)'T_i^{-1}(\boldsymbol{s}-F_i)+G_i)\Big)\Big\}d\boldsymbol{s}\\
&\quad +E_i\frac{\sqrt{|E_i||T_i|}}{\sqrt{|V_i||B_i|}}c\exp\Big\{-\frac{1}{2}G_i\Big\}\Phi(\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta,F_i,T_i)-E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\boldsymbol{1})'\\
&= c\exp\Big\{-\frac{1}{2}G_i\Big\}\frac{\sqrt{|E_i||T_i|}}{\sqrt{|V_i||B_i|}}\Phi(\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta,F_i,T_i)\Bigg(E_i+E_iV_i^{-1}\widetilde{X}_{1i}\beta(\widetilde{X}_{1i}\beta)'V_i^{-1}E_i\\
&\quad +E_iH'B^{-1}\Big(N+JJ'\Big)B_i^{-1}HE_i+EV_i^{-1}\widetilde{X}_{1i}\beta J'B^{-1}HE_i+E_iH'B^{-1}J(\widetilde{X}_{1i}\beta)'V_i^{-1}E_i\Bigg)\\
&\quad -E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\boldsymbol{1})E(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\boldsymbol{1})',
\end{aligned}$$

where J is the expected value of the truncated multivariate normal density, and N is the second moment of the latter density. More specifically,

$$\begin{aligned}
a &= \widetilde{\boldsymbol{X}}_{2i}\boldsymbol{\beta}+\boldsymbol{H}_i\widetilde{\boldsymbol{X}}_{1i}\boldsymbol{\beta}\\
J &= \boldsymbol{T}_i\big[-F_1(a_1)\quad -F_2(a_2)\quad \dots\quad -F_{\widetilde{p}_i}(a_{\widetilde{p}_i})\big]+\boldsymbol{F}_i,\\
N_{i,j} &= T_{i\,i,j}\sum_{k=1}^{\widetilde{p}_i}T_{i\,i,k}\frac{-T_{i\,j,k}a_kF_k(a_k)}{T_{i\,k,k}}+\sum_{k=1}^{\widetilde{p}_i}T_{i\,i,k}\sum_{q\neq k}\Big(T_{i\,j,q}-\frac{T_{i\,k,q}T_{i\,j,k}}{T_{i\,k,k}}\Big)\cdot -F_{k,q}(a_k,a_q)-J_iJ_k,\\
F_{k,q}(x,y) &= \int_{-\infty}^{a_1}\dots\int_{-\infty}^{a_{k-1}}\int_{-\infty}^{a_{k+1}}..\int_{-\infty}^{a_{q-1}}\int_{-\infty}^{a_{q+1}}\dots\int_{-\infty}^{a_{\widetilde{n}_i}}\phi(x,y,x_{-k-,-q},\boldsymbol{F}_i,\boldsymbol{T}_i)dx_{-k-,-q},\\
F_i(x_i) &= \int_{-\infty}^{a_1}\dots\int_{-\infty}^{o_{i-1}}\int_{-\infty}^{a_{i+1}}\dots\int_{-\infty}^{a_{\widetilde{p}_i}}\varphi(x_1,...x_{i-1},x,x_{i+1},...x_{\widetilde{p}_i})dx_{\widetilde{n}_i},..dx_{i+1}dx_{i-1}...dx_1\\
\varphi(x) &= \begin{cases}\frac{\phi(x,\boldsymbol{F}_i,\boldsymbol{T}_i)}{\Phi(\widetilde{\boldsymbol{X}}_{2i}\boldsymbol{\beta}+\boldsymbol{H}_i\widetilde{\boldsymbol{X}}_{1i}\boldsymbol{\beta},\boldsymbol{F}_i,\boldsymbol{T}_i)}, & \text{for } \boldsymbol{x}\leq\widetilde{\boldsymbol{X}}_{2i}\boldsymbol{\beta}+\boldsymbol{H}_i\widetilde{\boldsymbol{X}}_{1i}\boldsymbol{\beta},\\ 0, & \text{otherwise}\end{cases}
\end{aligned}$$

The above equations are deduced from Manjutah and Wilhelm (2012).

Agresti (2002) derived the distribution of transformed maximum likelihood parameters via the delta method

$$G(\hat{\beta})\rightarrow N\Big(\beta,\Big(\frac{\partial G(\beta)}{\partial\beta}\Big)'Var(\hat{\beta})\frac{\partial G(\beta)}{\partial\beta}\Big).$$

We will first sketch the derivative of the expected value with respect to a coefficient β_{12} of a predictor X_{12} of the continuous response and next sketch the derivation of a coefficient β_{22} of a predictor X_{22}

of the binary response.

$$\frac{\mathsf{E}(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\mathbf{1})}{\partial\beta_{12}} = \frac{1}{\Phi(\widetilde{X}_{2i}\beta, L^{-1})}\sqrt{\frac{|E_i||T_i|}{|V_i||B_i|}}\Big\{\Lambda\Phi(\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta, F_i, T_i)E_i(V_i^{-1}\widetilde{X}_{1i}\beta+H'B_i^{-1}F_i)$$
$$+e^{-0.5G}\lambda E_i(V_i^{-1}\widetilde{X}_{1i}\beta+H'B_i^{-1}F_i)$$
$$+e^{-0.5G}\Phi(\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta, F_i, T_i)\Big(EV_i^{-1}\widetilde{X}'_{12i}+E_iH'B_i^{-1}T_iB_i^{-1}HE_iV_i^{-1}\widetilde{X}'_{12i}\Big)$$
$$+E_iH'B_i^{-1}\cdot\Lambda\cdot T_i\Phi(o,T_i)\big[-F_1(o_1)\quad -F_2(o_2)\quad ...\quad -F_p(o_p)\big]+E_iH'B_i^{-1}e^{-0.5G}\nu\Big\},$$

with

$$\Lambda = \exp\Big\{-\frac{1}{2}G_i\Big\}\Big(-(\widetilde{X}'_{1ij}\beta)'V_i^{-1}\widetilde{X}'_{12i}+(\widetilde{X}'_{1ij}\beta)'V_i^{-1}E_iV_i^{-1}\widetilde{X}'_{12i}$$
$$+(\widetilde{X}'_{1ij}\beta)'V_i^{-1}E_iH'B_i^{-1}T_iB_i^{-1}HE_iV_i^{-1}\widetilde{X}'_{12i}\Big),$$
$$\lambda = \sum_{k=1}^{\widetilde{p}_i}(H\widetilde{X}'_{12i}-T_i\cdot B_i^{-1}HE_iV_i^{-1}\widetilde{X}_{12ij})_k\phi((\widetilde{X}'_{22ik}\beta+H\widetilde{X}_{1i}\beta)_k, F_{i,k}, T_{i,kk})$$
$$\times\Phi((\widetilde{X}'_{22i}\beta+H\widetilde{X}_{1i}\beta)_{-k}, F_{i,-k}, T_{i,-k-k}),$$
$$\nu = \sum_{k=1}^{\widetilde{p}_i}(H\widetilde{X}'_{12i}-T_i\cdot B_i^{-1}HE_iV_i^{-1}\widetilde{X}_{12ij})_k g_k(o_k)$$
$$g_k(x_k) = \int_{-\infty}^{o_1}\cdots\int_{-\infty}^{o_{i-1}}\int_{-\infty}^{o_{i+1}}\cdots\int_{-\infty}^{o_{\widetilde{p}_i}}[x_1..x_{k-1}o_kx_{k+1}..x_{\widetilde{p}_i}]'\phi([x_1..x_{k-1}o_kx_{k+1}..x_{\widetilde{p}_i}]', T_i)dx_{-k}.$$

Next, for a coefficient β_{22} of a predictor X_{22} of the binary response the derivative is the following

$$\frac{\mathsf{E}(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\mathbf{1})}{\partial\beta_{22}} = \frac{1}{\Phi^2(\widetilde{X}_{2i}\beta, L^{-1})}\sqrt{\frac{|E_i||T_i|}{|V_i||B_i|}}\exp\Big\{-\frac{1}{2}G_i\Big\} \tag{20}$$
$$\times\Big[E_i(V_i^{-1}\widetilde{X}_{1i}\beta+H'B_i^{-1}F_i)\Big(\Omega\Phi(\widetilde{X}_{2i}\beta, L^{-1})-\omega\Phi(\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta, F_i, T_i)\Big)$$
$$+E_iH'B_i^{-1}\Big(\xi\Phi(\widetilde{X}_{2i}\beta, L^{-1})-\omega\cdot T_i\Phi(o,T_i)\big[-F_1(o_1)\quad -F_2(o_2)\quad ...\quad -F_p(o_p)\big]\Big)\Big],$$

where

$$\omega = \sum_{k=1}^{\widetilde{p}}\widetilde{X}'_{22ik}\phi((\widetilde{X}_{2i}\beta)_k, L_{kk}^{-1})\Phi((\widetilde{X}_{2i}\beta)_{-k}, L_{-k,-k}^{-1})$$
$$\Omega = \sum_{k=1}^{\widetilde{p}}\widetilde{X}'_{22ik}\phi((\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta-F_i)_k, T_{kk})\Phi((\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta-F_i)_{-k},, T_{i\ -k,-k})$$
$$\xi = \sum_{k=1}^{\widetilde{p}_i}\widetilde{X}'_{22ik}g_k(o_k)$$

Next, the 95% prediction interval can be computed with the following general formula

$$\Bigg[E[\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}}|\widetilde{\boldsymbol{y}}_{\mathbf{2}\boldsymbol{i}} = \mathbf{1}] - 1.96\sqrt{E\left[\left(\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}} - E[\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}}|\widetilde{\boldsymbol{y}}_{\mathbf{2}\boldsymbol{i}} = \mathbf{1}]\right)^2\right] + \frac{\partial G(\beta)}{\partial \beta'} Var(\hat{\beta}) \frac{\partial G(\beta)}{\partial \beta}},$$
$$E[\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}}|\widetilde{\boldsymbol{y}}_{\mathbf{2}\boldsymbol{i}} = \mathbf{1}] + 1.96\sqrt{E\left[\left(\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}} - E[\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}}|\widetilde{\boldsymbol{y}}_{\mathbf{2}\boldsymbol{i}} = \mathbf{1}]\right)^2\right] + \frac{\partial G(\beta)}{\partial \beta'} Var(\hat{\beta}) \frac{\partial G(\beta)}{\partial \beta}}\Bigg].$$

C Conditional distribution of a subvector of the binary response given a subvector of the continuous response

C.1 Expected value

Let us derive (10), the conditional probability for the $\tilde{p}_i$-dimensional subvector of successes given a subvector of the continuous response $\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}$. Since the probability is equal to the conditional density, this is defined as

$$f(\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}} = \mathbf{1}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}) = \frac{\phi(\widetilde{X}_{1i}\boldsymbol{\beta}; V_i)\Phi(\widetilde{X}_{2i}\boldsymbol{\beta} - \alpha_i; B_i)}{\phi(\widetilde{X}_{1i}\boldsymbol{\beta}; V_i)} = \Phi(\widetilde{X}_{2i}\boldsymbol{\beta} - \alpha_i; B_i).$$

C.2 Confidence interval

In analogy with Appendix B.2, the gradients of the transformed parameters can be computed by the delta method in order to construct the confidence interval. First the standard error of a coefficient β_{12} of a predictor of the continuous response X_{12} is derived:

$$\frac{f(\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}} = \mathbf{1}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}})}{\partial\beta_{12}} = \sum_{k=1}^{\widetilde{p}_i} \gamma_{kk}^{-1/2} H_{ik} X_{12i} \phi(d_k) \Phi^{\tilde{R}_{B_i}^{(k)}}(\tilde{d}^{(k)}) \tag{21}$$

with

$$\begin{aligned}
\gamma &= \text{Diagonal matrix composed of the variances of } B_i \\
d &= \gamma^{-1/2}(X_{2i}\beta - H_i\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} + H_i X_{1i}\beta) \\
R_{B_i} &= \gamma^{-1/2} B_i \gamma^{-1/2} = (r_{ij})_{i,j=1}^{\widetilde{p}_i} \\
\tilde{d}^{(i)} &= \left(\frac{d_1 - r_{1,i}d_i}{\sqrt{1 - r_{1,i}^2}}, ..., \frac{d_{i-1} - r_{i-1,i}d_i}{\sqrt{1 - r_{i-1,i}^2}}, \frac{d_{i+1} - r_{i+1,i}d_i}{\sqrt{1 - r_{i+1,i}^2}}, ..., \frac{d_{\widetilde{p}_i} - r_{\widetilde{p}_i,i}d_i}{\sqrt{1 - r_{\widetilde{p}_i,i}^2}} \right).
\end{aligned}$$

The $(\widetilde{p}_i - 1) \times (\widetilde{p}_i - 1)$-dimensional correlation matrix $\tilde{R}_{B_i}^{(i)}$ has entries

$$\tilde{r}_{j,k}^{(i)} = \frac{r_{j',k'} - r_{j',i}r_{k',i}}{\sqrt{1 - r_{j',i}^2}\sqrt{1 - r_{k',i}^2}} \qquad i = 1...\tilde{p}_i; \quad j,k = 1.....\tilde{p}_i - 1$$

with

$$j' = \begin{cases} j & \text{if } j < i \\ j+1 & \text{if } j \geq i \end{cases} \qquad k' = \begin{cases} k & \text{if } k < i \\ k+1 & \text{if } k \geq i. \end{cases}$$

Next, the gradient of a coefficients β_{22} of one of the predictors of the binary response X_{22} is defined as

$$\frac{f(\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}} = \mathbf{1}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}})}{\partial\beta_{22}} = \sum_{k=1}^{\widetilde{p}_i} \gamma_{kk}^{-1/2} X_{22i} \phi(d_k) \Phi^{\tilde{R}_{B_i}^{(k)}} \tilde{d}^{(k)}, \tag{22}$$

where (21) and (22) have been derived from Prékopa (1995) p.204.

D Conditional distribution of a subvector of the continuous response given the binary response and a subvector of the continuous response

D.1 Expected value

Let us derive (11), the conditional expected value of the continuous subvector $\widetilde{\boldsymbol{Y}}^{a}_{1i}$=$(Y^{a_1}_{1i} Y^{a_2}_{1i} ... Y^{a_{n_a}}_{1i})$ given a distinct subvector of continuous responses $\widetilde{\boldsymbol{Y}}^{b}_{1i}$=$(Y^{b_1}_{1i} Y^{b_2}_{1i} ... Y^{b_{n_b}}_{1i})$ and the $\widetilde{p}_i$-dimensional binary subvector $\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}}$. In the following calculations, superscripts will indicate subvectors and submatrices. The superscript a denotes the rows a_1 until a_{n_a} and the superscript b denotes the rows b_1 until b_{n_b}. Analogously, the superscript bb denotes the submatrix with rows b_1 until b_{n_b} and columns b_1 until b_{n_b}. The superscript ab denotes the submatrix with rows a_1 until a_{n_a} and columns b_1 until b_{n_b}. Lastly, the superscript $.b$ denotes the columns b_1 until b_{n_b} from a matrix. The conditional expected value is equal to the integral over $\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}$ multiplied by the conditional distribution, which is defined as the joint distribution of $\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}} = [\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{Y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}]$ and $\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}}$ divided by the joint distribution of $\widetilde{\boldsymbol{Y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}$ and $\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}}$:

$$
\begin{aligned}
&E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} | \widetilde{\boldsymbol{Y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} = \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = 1] \\
&= \int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{1i}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{1i}=\infty} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \frac{\phi(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}; X_{1i}\beta; V_i)\Phi(X_{2i}\beta - H\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}} + HX_{1i}\beta; B_i)}{f(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}})} \Big\} d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\
&= \frac{1}{c} \int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{1i}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{1i}=\infty} \int_{\boldsymbol{t}=-\infty}^{\boldsymbol{t}=\widetilde{X}_{2i}\beta - H \begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(n_a+n_b+\widetilde{p}_i)}{2}}} \frac{1}{\sqrt{|V_i||B_i|}} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\
&\quad \text{exp}\Big\{ -\frac{1}{2}\Big((\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - \widetilde{X}_{1i}\beta)' V_i^{-1} (\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - \widetilde{X}_{1i}\beta) + \boldsymbol{t}' B_i^{-1} \boldsymbol{t} \Big) \Big\} d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} d\boldsymbol{t} \\
&= \frac{1}{c} \int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{1i}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{1i}=\infty} \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(n_a+n_b+\widetilde{p}_i)}{2}}} \frac{1}{\sqrt{|V_i||B_i|}} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\
&\quad \text{exp}\Big\{ -\frac{1}{2}\Big((\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - \widetilde{X}_{1i}\beta)' V_i^{-1} (\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - \widetilde{X}_{1i}\beta) + (\boldsymbol{s} - H \begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix})' B_i^{-1} (\boldsymbol{s} - H \begin{bmatrix} \widetilde{y}^{a}_{1i} \\ \widetilde{y}^{b}_{1i} \end{bmatrix}) \Big) \Big\} d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} d\boldsymbol{s} \\
&= \frac{1}{c} \int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{1i}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{1i}=\infty} \int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(n_a+n_b+\widetilde{p}_i)}{2}}} \frac{1}{\sqrt{|V_i||B_i|}} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\
&\quad \text{exp}\Big\{ -\frac{1}{2}\Big((\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - u_i)' \boldsymbol{E}_i^{-1} (\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - u_i) + O_i \Big) \Big\} d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} d\boldsymbol{s},
\end{aligned}
$$

where

$$
\begin{aligned}
c &= f(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}), \text{the marginal joint distribution} \\
\boldsymbol{s} &= \boldsymbol{t} - H\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}.
\end{aligned}
$$

Further,

$$
(\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - \widetilde{X}_{1i}\beta)' V_i^{-1} (\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - \widetilde{X}_{1i}\beta) + (\boldsymbol{s} - H \begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix})' B_i^{-1} (\boldsymbol{s} - H \begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix}) = (\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - u_i)' \boldsymbol{E}_i^{-1} (\begin{bmatrix} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} \end{bmatrix} - u_i) + O_i,
$$

with

$$
\begin{aligned}
\boldsymbol{E}_i^{-1} &= H'B_i^{-1}H + V_i^{-1} \\
l_i' &= -s'B_i^{-1}H - \widetilde{X}_{1i}\beta_1'V_i^{-1} \\
O_i &= s'B_i^{-1}s + \widetilde{X}_{1i}\beta_1'V_i^{-1}\widetilde{X}_{1i}\beta_1 - (-H'B_i^{-1}s - V_i^{-1}\widetilde{X}_{1i}\beta_1)'\boldsymbol{E}_i(-H'B_i^{-1}s - V_i^{-1}\widetilde{X}_{1i}\beta_1) \\
u_i &= -\boldsymbol{E}_i l_i.
\end{aligned}
$$

Integrating over $\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}$ results in

$$
\begin{aligned}
&E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}|\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} = \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = 1] \\
&= \frac{1}{c}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{\widetilde{p}_i}{2}}} \frac{\sqrt{|\boldsymbol{E}_i|}}{\sqrt{|V_i||B_i|}} \Big((\boldsymbol{u}_i^a + \boldsymbol{E}_i^{ab}\boldsymbol{E}_i^{bb-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - \boldsymbol{u}_i^b))\phi(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \boldsymbol{u}_i^b, \boldsymbol{E}_i^{bb})\Big)\text{exp}\Big\{-\frac{1}{2}O_i\Big\}d\boldsymbol{s} \\
&= \frac{1}{c}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(\widetilde{p}_i+n_b)}{2}}} \frac{\sqrt{|\boldsymbol{E}_i|}}{\sqrt{|V_i||B_i||\boldsymbol{E}_i^{bb}|}} (\boldsymbol{u}_i^a + \boldsymbol{E}_i^{ab}\boldsymbol{E}_i^{bb-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - \boldsymbol{u}_i^b)) \\
&\quad \text{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s}-F_i)'\boldsymbol{T}_i^{-1}(\boldsymbol{s}-F_i)+G_i\Big)\Big\}d\boldsymbol{s}
\end{aligned}
$$

where we have subsituted $O_i + (\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - \boldsymbol{u}_i^b)'(\boldsymbol{E}_i^{bb})^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - \boldsymbol{u}_i^b) = (\boldsymbol{s} - F_i)'T_i^{-1}(\boldsymbol{s} - F_i) + G_i$,

with

$$
\begin{aligned}
T_i^{-1} &= (\boldsymbol{E}_i H_i' B_i^{-1})^{b\prime}(\boldsymbol{E}_i^{bb}))^{-1}(\boldsymbol{E}_i H_i' B_i^{-1})^b + B_i^{-1} - (H_i'B_i^{-1})'\boldsymbol{E}_i(H_i'B_i^{-1}) \\
F_i &= T_i \cdot \Big((\boldsymbol{E}_i H_i' B_i^{-1})^{b\prime}(\boldsymbol{E}_i^{bb}))^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - (\boldsymbol{E}_i V^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}_1)^b) + (H_i'B_i^{-1})'\boldsymbol{E}_i(V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta})\Big) \\
G_i &= \Big(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - (\boldsymbol{E}_i V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}_1)^b\Big)'(\boldsymbol{E}_i^{bb}))^{-1}\Big(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - (\boldsymbol{E}_i V^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}_1)^b\Big) - F_i'T_i^{-1}F_i + \\
&\quad (\widetilde{X}_{1i}\boldsymbol{\beta})'V_i^{-1}(\widetilde{X}_{1i}\boldsymbol{\beta}) - (V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta})'\boldsymbol{E}_i(V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}).
\end{aligned}
$$

Integrating over $\boldsymbol{s}$ results in

$$
\begin{aligned}
E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} = \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = 1] &= \frac{e^{-0.5G_i}}{c(2\pi)^{\frac{n_b}{2}}} \frac{\sqrt{|\boldsymbol{E}_i||T_i|}}{\sqrt{|V_i||B_i||E_i^{bb}|}} \Phi(\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta, F_i, T_i) \\
&\quad \Big\{\Big((\boldsymbol{E}_i V^{-1}\widetilde{X}_{1i}\beta_1)^a + E^{ab}(E^{bb})^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - (\boldsymbol{E}_i V^{-1}\widetilde{X}_{1i}\beta_1)^b)\Big) \\
&+ \Big((EH'B^{-1})^a - E^{ab}(E^{bb})^{-1}(\boldsymbol{E}_i H B_i^{-1})^b\Big) \\
&\times \Big(T_i\big[-F_1(o_1) \quad -F_2(o_2) \quad \ldots \quad -F_p(o_p)\big] + F_i\Big)\Big\},
\end{aligned}
$$

where o,$F_i(x_i)$ and $\varphi(x_i)$ are defined in (18).

(11) is conditional on a sequence of successes (1,1,...,1)'. The expected value conditional on a arbitrary vector of failures and successes can be derived with appropriate contrasts. Given an arbitrary vector of failures and successes, the conditional expected value equals

$$
\begin{aligned}
E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{Y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} = \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}] &= \frac{e^{-0.5G_i}}{c(2\pi)^{\frac{n_b}{2}}} \frac{\sqrt{|\boldsymbol{E}_i||T_i|}}{\sqrt{|V_i||B_i||E_i^{bb}|}} \Phi(\widetilde{X}_{2i}\beta^* + H\widetilde{X}_{1i}\beta, F_i, T_i) \\
& \Bigg\{ \Bigg((\boldsymbol{E}_i V^{-1}\widetilde{X}_{1i}\beta_1)^a + E^{ab}(E^{bb})^{-1}(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - (\boldsymbol{E}_i V^{-1}\widetilde{X}_{1i}\beta_1)^b) \Bigg) \\
&+ \Bigg((EH'B^{-1})^a - E^{ab}(E^{bb})^{-1}(\boldsymbol{E}_i H' B_i^{-1})^b \Bigg) \\
&\times \Bigg(T_i\big[-F_1(o_1) \quad -F_2(o_2) \quad \ldots \quad -F_p(o_p)\big] + F_i) \Bigg) \Bigg\},
\end{aligned}
$$

with

$$
\begin{aligned}
\widetilde{X}_{2i}\beta^* &= s \diamond \widetilde{X}_{2i}\beta \\
\widetilde{Z}^*_{2i} &= J \diamond Z_{2i},
\end{aligned}
$$

where $\diamond$ indicates pairwise multiplication and s and J are defined as in (16) and (17) respectively.

D.2 Prediction interval

The prediction interval of a subvector of continuous responses conditional on another subvector of continuous responses and a subvector of binary responses is defined analogously to Appendix B.2. We will first derive the second central moment of the conditional distribution and next derive the standard errors of the transformed parameters.

The second central moment of (11) equals

$$
\begin{aligned}
& E\Big[\big(\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\big)\big(\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\big)'\Big] \\
= \; & \int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=\infty} \Big[\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\Big]\Big[\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\Big]' \cdot f(\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = 1) d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\
= \; & \int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=+\infty} \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}' f(\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = 1) \\
& - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]' f(\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = 1) \\
& - \Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{a}\,\prime}_{\boldsymbol{1i}} E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}] + E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\widetilde{\boldsymbol{y}}^{\boldsymbol{a}\,\prime}_{\boldsymbol{1i}}\Big) \cdot f(\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = 1) d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} \\
= \; & \int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=+\infty} \widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}\,\prime}_{\boldsymbol{1i}} f(\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = 1) d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]' \\
= \; & \frac{1}{c}\int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=\infty}\int_{\boldsymbol{t}=-\infty}^{\boldsymbol{t}=\widetilde{X}_{2i}\beta - H\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(n_a+n_b+\widetilde{p}_i)}{2}}}\frac{1}{\sqrt{|V_i||B_i|}}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}\,\prime}_{\boldsymbol{1i}} \\
& \times\exp\Big\{-\frac{1}{2}\Big(\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - \widetilde{X}_{1i}\beta\big)'V_i^{-1}\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - \widetilde{X}_{1i}\beta\big) + \boldsymbol{t}'B_i^{-1}\boldsymbol{t}\Big)\Big\} d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} d\boldsymbol{t} \\
& - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]' \\
= \; & \frac{1}{c}\int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=\infty}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(n_a+n_b+\widetilde{p}_i)}{2}}}\frac{1}{\sqrt{|V_i||B_i|}}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}\,\prime}_{\boldsymbol{1i}} \\
& \times\exp\Big\{-\frac{1}{2}\Big(\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - \widetilde{X}_{1i}\beta\big)'V_i^{-1}\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - \widetilde{X}_{1i}\beta\big) + \big(\boldsymbol{s} - H\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix}\big)'B_i^{-1}\big(\boldsymbol{s} - H\begin{bmatrix}\widetilde{y}^{a}_{1i}\\ \widetilde{y}^{b}_{1i}\end{bmatrix}\big)\Big)\Big\} d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} d\boldsymbol{s} \\
& - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]' \\
= \; & \frac{1}{c}\int_{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=-\infty}^{\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}=\infty}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{(n_a+n_b+\widetilde{p}_i)}{2}}}\frac{1}{\sqrt{|V_i||B_i|}}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}\,\prime}_{\boldsymbol{1i}} \\
& \exp\Big\{-\frac{1}{2}\Big(\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - u_i\big)'E_i^{-1}\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - u_i\big) + O_i\Big)\Big\} d\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} d\boldsymbol{s} - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]',
\end{aligned}
$$

where

$$
\begin{aligned}
c &= f(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}), \text{the marginal joint distribution} \\
\boldsymbol{s} &= \boldsymbol{t} - H\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}.
\end{aligned}
$$

Further,

$$
\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - \widetilde{X}_{1i}\beta\big)'V_i^{-1}\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - \widetilde{X}_{1i}\beta\big) + \big(\boldsymbol{s} - H\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix}\big)'B_i^{-1}\big(\boldsymbol{s} - H\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix}\big) = \big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - u_i\big)'E_i^{-1}\big(\begin{bmatrix}\widetilde{\boldsymbol{y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}\\ \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}\end{bmatrix} - u_i\big) + O_i,
$$

with

$$
\begin{aligned}
E_i^{-1} &= H'B_i^{-1}H + V_i^{-1} \\
l_i' &= -s'B_i^{-1}H - \widetilde{X}_{1i}\beta_1'V_i^{-1} \\
O_i &= s'B_i^{-1}s + \widetilde{X}_{1i}\beta_1'V_i^{-1}\widetilde{X}_{1i}\beta_1 - (-H'B_i^{-1}s - V_i^{-1}\widetilde{X}_{1i}\beta_1)'E(-H'B_i^{-1}s - V_i^{-1}\widetilde{X}_{1i}\beta_1) \\
u_i &= -E_i l_i.
\end{aligned}
$$

Applying the formulas of the conditional mean and variance in the multivariate normal distribution, integration over $\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}$ results in

$$
\begin{aligned}
&E\Big[\big(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}} - E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \mathbf{1}]\big)\big(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}} - E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \mathbf{1}]\big)'\Big] \\
&= \frac{1}{c}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{\widetilde{p}}{2}}}\frac{\sqrt{|E_i|}}{\sqrt{|V_i||B_i|}}\Big\{E_i^{aa} - E_i^{ab}(E_i^{bb})^{-1}E_i^{ba} + \Big(u_i^a + E_i^{ab}(E_i^{bb})^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - u_i^b)\Big) \\
&\quad \times\Big(u_i^a + E_i^{ab}(E_i^{bb})^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - u_i^b)\Big)'\Big\}\phi(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, u_i^b, E_i^{bb})\mathsf{exp}\{-\frac{1}{2}O_i\}d\boldsymbol{s} \\
&\quad -E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \mathbf{1}]E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \mathbf{1}]' \\
&= \frac{1}{c}\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta} \frac{1}{(2\pi)^{\frac{n_b+\widetilde{p}}{2}}}\frac{\sqrt{|E_i|}}{\sqrt{|V_i||B_i|}}\Big\{E_i^{aa} - E_i^{ab}(E_i^{bb})^{-1}E_i^{ba} + \Big(u_i^a + E_i^{ab}(E_i^{bb})^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - u_i^b)\Big) \\
&\quad \times\Big(u_i^a + E_i^{ab}(E_i^{bb})^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - u_i^b)\Big)'\Big\}\mathsf{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s} - F_i)'T_i^{-1}(\boldsymbol{s} - F_i) + G_i\Big)\Big\}d\boldsymbol{s} \\
&\quad -E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \mathbf{1}]E[\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}}|\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \mathbf{1}]'.
\end{aligned}
$$

Here we have rewritten $\phi(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, u_i^b, E_i^{bb})\mathsf{exp}\{-\frac{1}{2}O_i\} = \mathsf{exp}\Big\{-\frac{1}{2}\Big((\boldsymbol{s} - F_i)'T_i^{-1}(\boldsymbol{s} - F_i) + G_i\Big)$, with

$$
\begin{aligned}
T_i^{-1} &= (E_iH'B_i^{-1})'^b(E_i^{bb})^{-1}(E_iH'B_i^{-1})^b + B_i^{-1} - (H'B_i^{-1})'E_i(H'B_i^{-1}) \\
F_i &= T_i \cdot \Big((E_iH'B_i^{-1})'^b(E_i^{bb})^{-1}(\widetilde{y}_{1i}^b - (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^b) + (H'B_i^{-1})'E_i(V_i^{-1}\widetilde{X}_{1i}\beta)\Big) \\
G_i &= \Big(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - (E_iV_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}_1)^b\Big)'(E_i^{bb})^{-1}\Big(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - (E_iV^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}_1)^b\Big) - F_i'T_i^{-1}F_i + \\
&\quad (\widetilde{X}_{1i}\boldsymbol{\beta})'V_i^{-1}(\widetilde{X}_{1i}\boldsymbol{\beta}) - (V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta})'E_i(V_i^{-1}\widetilde{X}_{1i}\boldsymbol{\beta}).
\end{aligned}
$$

Integrating over s results in

$$
\begin{aligned}
&E\Big[\big(\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}} - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]\big)^2\Big] \\
&= \nu\big(E_i^{aa} - E_i^{ab}(E_i^{bb})^{-1}E_i^{ba}\big) + \nu\Bigg\{(E_iH'B_i^{-1})^a\Big(N + JJ'\Big)(E_iH'B_i^{-1})^{a\prime} \\
&\quad +(E_iH'B_i^{-1})^a J((\widetilde{X}_{1i}\beta_1)'V_i^{-1}E_i)^a + (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^a J')(E_iH'B_i^{-1})^{a\prime} + (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^a(E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^{a\prime}\Bigg\} \\
&\quad +\nu\Bigg\{(E_iH'B_i^{-1})^a J\Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - ((\widetilde{X}_{1i}\beta_1)'V_i^{-1}E_i)^b\Big) - (E_iH'B_i^{-1})^a\Big(N + JJ'\Big)(E_iH'B_i^{-1})^{b\prime} \\
&\quad +(E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^a\Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - ((\widetilde{X}_{1i}\beta_1)'V_i^{-1}E_i)^b\Big)' - (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^a J'(E_iH'B_i^{-1})^{b\prime}\Bigg\}(E_i^{bb})^{-1}E_i^{ba} \\
&\quad +\nu E_i^{ab}(E_i^{bb})^{-1}\Bigg\{\Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^b\Big)((\widetilde{X}_{1i}\beta_1)'V_iE_i)^a - (E_iH'B_i^{-1})^b J((\widetilde{X}_{1i}\beta_1)'V_iE_i)^a \\
&\quad +\Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^b\Big)J'(E_iH'B_i^{-1})^{a\prime} - (E_iH'B_i^{-1})^b\Big(N + JJ'\Big)(E_iH'B_i^{-1})^{a\prime}\Bigg\} \\
&\quad +\nu E_i^{ab}(E_i^{bb})^{-1}\Bigg\{(E_iH'B_i^{-1})^b\Big(N + JJ')\Big)(E_iH'B_i^{-1})^{b\prime} - (E_iH'B_i^{-1})^b J\Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^b\Big)' \\
&\quad -\Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^b\Big)J'(E_iH'B_i^{-1})^{b\prime} + \Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^b\Big)\Big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} - (E_iV_i^{-1}\widetilde{X}_{1i}\beta_1)^b)\Big)'\Bigg\} \\
&\quad \times(E_i^{bb})^{-1}E_i^{ba} - E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]E[\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1}]',
\end{aligned}
$$

with J as the expected value of the truncated multivariate normal density, and N is the second central moment of the latter density. These are defined in (19).

The standard errors of the transformed parameters are derived analogously to Appendix B.2 with the delta method. The derivative of the expected value with respect to β_{12}, and arbitrary coefficient of a predictor of the continuous response X_{12} is the following

$$
\begin{aligned}
\frac{\mathsf{E}(\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{1i}}|\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}},\widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}=\boldsymbol{1})}{\partial\beta_{12}} &= \frac{1}{c^2(2\pi)^{\frac{n_b}{2}}}\frac{\sqrt{|E_i||T_i|}}{\sqrt{|V_i||B_i||E_i^{bb}|}}\Bigg[c\Lambda\Phi(\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta,F_i,T_i) \\
&\times \Bigg\{\Big((EV^{-1}\widetilde{X}_{1i}\beta_1)^a+E_i^{ab}(E_i^{bb})^{-1}\big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}-(EV^{-1}\widetilde{X}_{1i}\beta_1)^b\big)\Big) \\
&+ \Big((E_iH'B_i^{-1})^a-E^{ab}(E^{bb})^{-1}(E_iH'B_i^{-1})^b\Big) \\
&\quad \Big(T_i\big[-F_1(a_1)\quad -F_2(a_2)\quad \ldots\quad -F_p(a_p)\big]+F_i)\Big)\Bigg\} \\
&+ ce^{-0.5G}\Theta\Bigg\{\Big((EV^{-1}\widetilde{X}_{1i}\beta_1)^a+E_i^{ab}(E_i^{bb})^{-1}\big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}-(EV^{-1}\widetilde{X}_{1i}\beta_1)^b\big)\Big) \\
&+ \Big((E_iH'B_i^{-1})^a-E^{ab}(E^{bb})^{-1}(E_iH'B_i^{-1})^b\Big)F_i\Bigg\} \\
&+ ce^{-0.5G}\Phi(\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta,F_i,T_i)\Bigg\{(E_iV_i^{-1})^a\widetilde{X}_{12i}-E_i^{ab}(E_i^{bb})^{-1}(E_iV^{-1})^b\widetilde{X}_{12i}\Bigg\} \\
&+ ce^{-0.5G}\Bigg\{\Big((E_iH'B_i^{-1})^a-E^{ab}(E^{bb})^{-1}(E_iH'B_i^{-1})^b\Big)\nu\Bigg\} \\
&- \lambda e^{-0.5G}\Phi(\widetilde{X}_{2i}\beta+H\widetilde{X}_{1i}\beta,F_i,T_i) \\
&\times \Bigg\{\Big((EV^{-1}\widetilde{X}_{1i}\beta_1)^a+E_i^{ab}(E_i^{bb})^{-1}\big(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}-(EV^{-1}\widetilde{X}_{1i}\beta_1)^b\big)\Big) \\
&+ \Big((E_iH'B_i^{-1})^a-E^{ab}(E^{bb})^{-1}(E_iH'B_i^{-1})^b\Big) \\
&\quad \Big(T_i\big[-F_1(a_1)\quad -F_2(a_2)\quad \ldots\quad -F_p(a_p)\big]+F_i)\Big)\Bigg\}\Bigg],
\end{aligned}
\tag{23}
$$

with

$$
\begin{aligned}
\Lambda &= -e^{-\frac{1}{2}G}\Bigg\{(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}-(EV^{-1}\widetilde{X}_{1i}\beta_1)^b)'(E^{bb})^{-1}(EV^{-1}\widetilde{X}_{12i})^b) \\
&+ \widetilde{X}_{1i}\beta_1'V^{-1}\widetilde{X}_{12i}-(V^{-1}\widetilde{X}_{1i}\beta_1)'E_iV^{-1}\widetilde{X}_{12i}-F_iTi^{-1}\delta_i'\Bigg\} \\
\Theta &= \sum_{k=1}^{\widetilde{p}}\Gamma_{kk}^{-1/2}\big(H_kX_{12i}-F_{ik}'\big)\phi\big[(X_{2i}\beta+HX_{1i}\beta-F_i)_k,T_{kk}\big]\Phi^{\tilde{R}_{T_i}^{(k)}}(\tilde{z}^{(k)}),
\end{aligned}
$$

with

$$
\begin{aligned}
\delta_i' &= T_i\bigg(-(E_i H' B^{-1})^{b\prime}(E_i^{bb})^{-1}(E_i V^{-1})^b \widetilde{X}_{12i} + (H' B_i^{-1})' E_i (V_i^{-1}\widetilde{X}_{12i})\bigg) \\
\Gamma &= \text{Diagonal matrix composed of the variances of } T_i \\
z &= \Gamma^{-1/2}(X_{2i}\beta + H X_{1i}\beta - F_i) \\
R_{T_i} &= \Gamma^{-1/2} T_i \Gamma^{-1/2} = (r_{ij})_{i,j=1}^{\widetilde{p}_i} \\
\tilde{z}^{(i)} &= \left(\frac{z_1 - r_{1,i}z_i}{\sqrt{1-r_{1,i}^2}}, ..., \frac{z_{i-1} - r_{i-1,i}z_i}{\sqrt{1-r_{i-1,i}^2}}, \frac{z_{i+1} - r_{i+1,i}z_i}{\sqrt{1-r_{i+1,i}^2}}, ..., \frac{z_{\widetilde{p}_i} - r_{\widetilde{p}_i,i}z_i}{\sqrt{1-r_{\widetilde{p}_i,i}^2}}\right).
\end{aligned}
$$

The $(\widetilde{p}_i - 1) \times (\widetilde{p}_i - 1)$-dimensional correlation matrix $\tilde{R}_{T_i}^{(i)}$ has entries

$$
\tilde{r}_{j,k}^{(i)} = \frac{r_{j',k'} - r_{j',i}r_{k',i}}{\sqrt{1-r_{j',i}^2}\sqrt{1-r_{k',i}^2}} \qquad i = 1...\tilde{p}_i; \quad j,k = 1.....\tilde{p}_i - 1
$$

with

$$
j' = \begin{cases} j & \text{if } j < i \\ j+1 & \text{if } j \geq i \end{cases} \qquad k' = \begin{cases} k & \text{if } k < i \\ k+1 & \text{if } k \geq i. \end{cases}
$$

Further,

$$
\begin{aligned}
\boldsymbol{\nu} &= \sum_{k=1}^{\widetilde{p}_i} (\boldsymbol{H}_i \widetilde{\boldsymbol{X}}_{12i} + \boldsymbol{\delta}_i)_k g_k(a_k) + \Phi(\widetilde{\boldsymbol{X}}_{2i}\boldsymbol{\beta} + \boldsymbol{H}_i \widetilde{\boldsymbol{X}}_{1i}\boldsymbol{\beta}, \boldsymbol{F}_i, \boldsymbol{T}_i)\boldsymbol{\delta}_{ik}, \\
a &= \widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta, \\
g_k(x_k) &= \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_{i-1}} \int_{-\infty}^{a_{i+1}} \cdots \int_{-\infty}^{a_{\widetilde{p}_i}} [x_1..x_{k-1}a_k x_{k+1}..x_{\widetilde{p}_i}]' \phi([x_1..x_{k-1}a_k x_{k+1}..x_{\widetilde{p}_i}]', T_i) dx_{-k}.
\end{aligned}
$$

Next,

$$
\begin{aligned}
\lambda &= \sum_{k=1}^{\widetilde{p}} \gamma_{kk}^{-1/2} H_k^* X_{12i}^b \phi\big[(X_{2i}\beta - H_i^* \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} + H_i^* X_{1i}^b \beta)_k, B_{kk}\big] \Phi^{\tilde{R}_{B_i^*}^{(k)}}(\tilde{d}^{(k)}) \phi(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, X_{1i}^b\beta, V_i^{bb}) \quad (24) \\
&\quad + \Phi(X_{2i}\beta - H^* \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} + H^* X_{1i}^b \beta, B_i^*)(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - X_{1i}^b \beta)'(V_i^{bb})^{-1} X_{12i}^b \phi(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, X_{1i}^b\beta, V_i^{bb})
\end{aligned}
$$

where

$$\begin{aligned}
B_i^{*-1} &= I - Z_{2i}K_i^* Z_{2i}', \\
K_i^{*-1} &= D^{-1} + Z^{b\prime}{}_{1i}(\Sigma_i^{bb})^{-1} Z_{1i}^b + Z_{2i}' Z_{2i}, \\
H_i^* &= -B_i^* Z_{2i} K_i^* Z^{b\prime}{}_{1i}(\Sigma_i^{bb})^{-1}. \\
\gamma &= \text{Diagonal matrix composed of the variances of } B_i^* \\
d &= \gamma^{-1/2}(X_{2i}\beta - H_i^* \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} + H_i^* X_{1i}^b \beta) \\
R_{B_i^*} &= \gamma^{-1/2} B_i^* \gamma^{-1/2} = (r_{ij})_{i,j=1}^{\widetilde{p}_i} \\
\tilde{d}^{(i)} &= \left(\frac{d_1 - r_{1,i}d_i}{\sqrt{1 - r_{1,i}^2}}, ..., \frac{d_{i-1} - r_{i-1,i}d_i}{\sqrt{1 - r_{i-1,i}^2}}, \frac{d_{i+1} - r_{i+1,i}d_i}{\sqrt{1 - r_{i+1,i}^2}}, ..., \frac{d_{\widetilde{p}_i} - r_{\widetilde{p}_i,i}d_i}{\sqrt{1 - r_{\widetilde{p}_i,i}^2}} \right).
\end{aligned}$$

The $(\widetilde{p}_i - 1) \times (\widetilde{p}_i - 1)$-dimensional correlation matrix $\tilde{R}_{B_i^*}^{(i)}$ has entries

$$\tilde{r}_{j,k}^{(i)} = \frac{r_{j',k'} - r_{j',i} r_{k',i}}{\sqrt{1 - r_{j',i}^2}\sqrt{1 - r_{k',i}^2}} \qquad i = 1...\tilde{p}_i; \quad j,k = 1.....\tilde{p}_i - 1$$

with

$$j' = \begin{cases} j & \text{if } j < i \\ j+1 & \text{if } j \geq i \end{cases} \qquad k' = \begin{cases} k & \text{if } k < i \\ k+1 & \text{if } k \geq i. \end{cases}$$

The derivative of the expected value with respect to a coefficient β_{22} of a predictor X_{22} of the binary response

$$\begin{aligned}
\frac{\mathsf{E}(\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}^{\boldsymbol{a}} | \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}} = \boldsymbol{1})}{\partial \beta_{22}} &= \frac{e^{-0.5G_i}}{c^2 (2\pi)^{\frac{n_b}{2}}} \frac{\sqrt{|E_i||T_i|}}{\sqrt{|V_i||B_i||E_i^{bb}|}} \Bigg[c\omega \Bigg\{ \Big((E_i V^{-1} \widetilde{X}_{1i}\beta_1)^a + \\
& E_i^{ab}(E_i^{bb})^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - (E_i V^{-1}\widetilde{X}_{1i}\beta_1)^b) \Big) + \Big((E_i H' B_i^{-1})^a - E_i^{ab}(E_i^{bb})^{-1}(E_i H' B_i^{-1})^b \Big) \\
&\times \Big(T_i \big[-F_1(a_1) \quad -F_2(a_2) \quad ... \quad -F_p(a_p) \big] + F_i \Big) \Big) \Bigg\} \\
&+ c \Bigg\{ \Big((E_i H' B_i^{-1})^a - E_i^{ab}(E_i^{bb})^{-1}(E_i H' B_i^{-1})^b \Big) \zeta \Bigg\} \\
&- \Omega \Phi(\widetilde{X}_{2i}\beta + H\widetilde{X}_{1i}\beta, F_i, T_i) \Bigg\{ \Big((E_i V^{-1}\widetilde{X}_{1i}\beta_1)^a + E_i^{ab}(E_i^{bb})^{-1}(\widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}^{\boldsymbol{b}} - (E_i V^{-1}\widetilde{X}_{1i}\beta_1)^b) \Big) \\
&+ \Big((E_i H' B_i^{-1})^a - E_i^{ab}(E_i^{bb})^{-1}(E_i H' B_i^{-1})^b \Big) \\
& \Big(T_i \big[-F_1(a_1) \quad -F_2(a_2) \quad ... \quad -F_p(a_p) \big] + F_i \Big) \Big) \Bigg\} \Bigg],
\end{aligned} \tag{25}$$

with

$$
\begin{aligned}
\zeta &= \sum_{k=1}^{\widetilde{p}_i} \widetilde{X}'_{22ik} g_k(o_k) \\
\Omega &= \phi(\widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}}, X^b_{1i}\beta; V^{bb}_i) \sum_{k=1}^{\widetilde{p}_i} \gamma_{kk}^{-1/2} X_{22ik} \phi\big[(X_{2i}\beta - H^*_i \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{1i}} + H^*_i X^b_{1i}\beta)_k, B_{kk}\big] \Phi^{\tilde{R}^{(k)}_{B_i}}(\tilde{d}^{(k)}), \\
\omega &= \sum_{k=1}^{\widetilde{p}} \Gamma_{kk}^{-1/2} X_{22ik} \phi(z_{1_k}) \Phi^{\tilde{R}^{(k)}_{T_i}}(\tilde{z}^{(k)}_1),
\end{aligned}
$$

(24), (24), (26) and (26) have been derived from Prékopa (1995) p.204.

E Conditional distribution of a subvector of the binary response given a subvector of the continuous response and a subvector of the binary response

E.1 Expected value

Let us derive (12), the conditional probability for the p_a-dimensional subvector of successes $\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{2i}}$ given a subvector of the $\tilde{n}_i$-dimensional subvector of continuous responses $\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}$ and a p_b-dimensional subvector of successes $\widetilde{\boldsymbol{Y}}^{\boldsymbol{b}}_{\boldsymbol{2i}}$.In the following calculations, superscripts will indicate subvectors and submatrices. The superscript a denotes the rows a_1 until a_{p_a} and the superscript b denotes the rows b_1 until p_b. Analogously, the superscript bb denotes the submatrix with rows b_1 until p_b and columns b_1 until p_b. The superscript ab denotes the submatrix with rows a_1 until a_{p_a} and columns b_1 until p_b.

Since conditional probability is equal to the conditional density, the conditional probability equals the joint distribution of $\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}} = [\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{2i}}, \widetilde{\boldsymbol{Y}}^{\boldsymbol{b}}_{\boldsymbol{2i}}]$ and $\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}$ divided by the joint distribution of $\widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}$ and $\widetilde{\boldsymbol{Y}}^{\boldsymbol{b}}_{\boldsymbol{2i}}$.

$$
\begin{aligned}
f(\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{2i}} = \mathbf{1} | \widetilde{\boldsymbol{Y}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{2i}} = \mathbf{1}) &= \frac{\phi(\widetilde{X}_{1i}\boldsymbol{\beta}; V_i)\Phi(\widetilde{X}_{2i}\boldsymbol{\beta} - H_i(\widetilde{\mathbf{Y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B_i)}{\phi(\widetilde{X}_{1i}\boldsymbol{\beta}; V_i)\Phi(\widetilde{X}^b_{2i}\boldsymbol{\beta} - H^b_i(\widetilde{\mathbf{Y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B^{bb}_i)} \\
&= \frac{\Phi(\widetilde{X}_{2i}\boldsymbol{\beta} - H_i(\widetilde{\mathbf{Y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B_i)}{\Phi(\widetilde{X}^b_{2i}\boldsymbol{\beta} - H^b_i(\widetilde{\mathbf{Y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B^{bb}_i)},
\end{aligned}
$$

where $\widetilde{X}_{2i}\boldsymbol{\beta} = [\widetilde{X}^a_{2i}\boldsymbol{\beta}, \widetilde{X}^b_{2i}\boldsymbol{\beta}]$ and $\widetilde{p}_i = p_a + p_b$.

E.2 Confidence interval

In order to construct the confidence interval, the standard errors of the transformed parameters can be computed by the delta method. First the gradient of a coefficient β_{12} of a predictor of the continuous response X_{12} is derived:

$$
\begin{aligned}
\frac{f(\widetilde{\boldsymbol{Y}}^{\boldsymbol{a}}_{\boldsymbol{2i}} = \mathbf{1} | \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}^{\boldsymbol{b}}_{\boldsymbol{2i}} = \mathbf{1})}{\partial \beta_{12}} &= \Bigg(\sum_{k=1}^{\widetilde{p}_i} \gamma_{kk}^{-1/2} H_k X_{12i} \phi(d_k) \Phi^{\tilde{R}^{(k)}_{B_i}}(\tilde{d}^{(k)}) \Phi(\widetilde{X}^b_{2i}\boldsymbol{\beta} - H^b(\widetilde{\mathbf{Y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B^{bb}_i) \\
&- \sum_{k=1}^{\widetilde{p}^b_i} \gamma_{kk}^{-1/2} H^b_k X_{12i} \phi(d_k) \Phi^{\tilde{R}^{(k)}_{B^{bb}_i}}(\tilde{d}^{(k)}) \Phi(\widetilde{X}_{2i}\boldsymbol{\beta} - H(\widetilde{\mathbf{Y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B_i) \Bigg) \\
&\times \left(\Phi(\widetilde{X}^b_{2i}\boldsymbol{\beta} - H^b(\widetilde{\mathbf{Y}}_{\boldsymbol{1i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B^{bb}_i) \right)^{-2},
\end{aligned} \tag{26}
$$

where γ, d, R_{B_i}, $\tilde{d}^{(i)}$ and $\tilde{R}^{(i)}_{B_i}$ are defined in (21).

Next, the gradient of a coefficient β_{22} of one of the predictors of the binary response X_{22} is defined as

$$\frac{f(\widetilde{\boldsymbol{Y}}_{\boldsymbol{2i}}^{\boldsymbol{a}} = \mathbf{1} | \widetilde{\boldsymbol{y}}_{\boldsymbol{1i}}, \widetilde{\boldsymbol{y}}_{\boldsymbol{2i}}^{\boldsymbol{b}} = \mathbf{1})}{\partial \beta_{22}} = \Bigg(\sum_{k=1}^{\widetilde{p}_i} \gamma_{kk}^{-1/2} X_{22ik} \phi(d_k) \Phi^{\tilde{R}_{B_i}^{(k)}}(\tilde{d}^{(k)}) \Phi(\widetilde{X}_{2i}^b \boldsymbol{\beta} - H^b(\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B_i^{bb})$$

$$- \sum_{k=1}^{\widetilde{p}_i^b} \gamma_{kk}^{-1/2} X_{22ik}^b \phi(d_k) \Phi^{\tilde{R}_{B_i^{bb}}^{(k)}}(\tilde{d}^{(k)}) \Phi(\widetilde{X}_{2i}\boldsymbol{\beta} - H(\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B_i) \Bigg)$$

$$\times \left(\Phi(\widetilde{X}_{2i}^b \boldsymbol{\beta} - H^b(\widetilde{\mathbf{Y}}_{\mathbf{1}\boldsymbol{i}} - \widetilde{X}_{1i}\boldsymbol{\beta}); B_i^{bb}) \right)^{-2}, \tag{27}$$

where (26) and (27) have been derived from Prékopa (1995) p.204.

F Computation of the correlation of single responses

In this section (13) will be derived. Because the elements of $\boldsymbol{Y}_{1i}$ and $\boldsymbol{Y}_{2i}$ are independent conditional on the random effects, we have for the expected value of the product:

$$
\begin{aligned}
&E[Y_{1ij}Y_{2ik}] \\
&= E_{\boldsymbol{b}_i}[E(Y_{1ij}, Y_{2ik}|\boldsymbol{b}_i)] \\
&= E_{\boldsymbol{b}_i}[E(Y_{1ij}|\boldsymbol{b}_i)E(Y_{2ik}|\boldsymbol{b}_i)] \\
&= E_{\boldsymbol{b}_i}[(X'_{1ij}\beta + Z'_{1ij}\boldsymbol{b}_i)\cdot \Phi(X'_{2ij}\beta + Z_{2ij}\boldsymbol{b}_i)] \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \int_{t=-\infty}^{t=X'_{2ik}\beta+Z'_{2ik}\boldsymbol{b}_i} \frac{1}{(2\pi)^{q/2}|D|^{1/2}\sqrt{2\pi}}\text{exp}\Big\{-\frac{1}{2}(\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i + t^2)\Big\}(X'_{1ij}\beta + Z'_{1ij}\boldsymbol{b}_i)d\boldsymbol{b}_idt \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \int_{s=-\infty}^{s=X'_{2ik}\beta} \frac{1}{(2\pi)^{q/2}|D|^{1/2}\sqrt{2\pi}}\text{exp}\Big\{-\frac{1}{2}(\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i + (s+z'_{2ik}\boldsymbol{b}_i)^2)\Big\}(X'_{1ij}\beta + Z'_{1ij}\boldsymbol{b}_i)d\boldsymbol{b}_ids \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \int_{s=-\infty}^{s=X'_{2ik}\beta} \frac{1}{(2\pi)^{q/2}|D|^{1/2}\sqrt{2\pi}}\text{exp}\Big\{-\frac{1}{2}\Big(\frac{s}{(1-(Z'_{2ik}(D^{-1}+Z_{2ik}Z'_{2ik})^{-1}Z_{2ik}))^{-1/2}}\Big)^2\Big\} \\
&\quad \times \text{exp}\Big\{-\frac{1}{2}((\boldsymbol{b}_i-k)'[(D^{-1}+Z_{2ik}Z'_{2ik})^{-1}]^{-1}(\boldsymbol{b}_i-k)\Big\}(X'_{1ij}\beta + Z'_{1ij}\boldsymbol{b}_i)d\boldsymbol{b}_ids.
\end{aligned}
\tag{28}
$$

Here, we have rewritten the variable $t = s+z'_{2ik}\boldsymbol{b}_i$ and $\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i+t^2=(\boldsymbol{b}_i-k)'[(D^{-1}+Z_{2ik}Z'_{2ik})^{-1}]^{-1}(\boldsymbol{b}_i-k) + \frac{s}{(1-(Z'_{2ik}(D^{-1}+Z_{2ik}Z'_{2ik})^{-1}Z_{2ik}))^{-1/2}})^2$, where

$$k = -(D^{-1} + Z'_{2ik}Z_{2ik})^{-1}Z'_{2ik}s$$

Next, we integrate over the random effects, which results in

$$
\begin{aligned}
&E[Y_{1ij}Y_{2ik}] \\
&= \int_{s=-\infty}^{s=X'_{2ik}\beta} \text{exp}\Big\{-\frac{1}{2}\Big(\frac{s}{(1-(Z'_{2ik}(D^{-1}+Z_{2ik}Z'_{2ik})^{-1}Z_{2ik}))^{-1/2}}\Big)^2\Big\} \\
&\quad \times\Bigg(\frac{1}{|D|^{1/2}\sqrt{2\pi}}X'_{1ij}\beta\frac{1}{|D^{-1}+Z_{2ik}Z'_{2ik}|^{1/2}} + Z'_{1ij}\frac{1}{|D|^{1/2}\sqrt{2\pi}}\frac{1}{|D^{-1}+Z_{2ik}Z'_{2ik}|^{1/2}}k\Bigg)ds \\
&= \int_{u=-\infty}^{u=\frac{X'_{2ik}\beta}{L^{-1/2}}} L^{-1/2}\text{exp}\Big\{-\frac{1}{2}u^2\Big\}\Big(\frac{1}{|D|^{1/2}\sqrt{2\pi}}X'_{1ij}\beta\frac{1}{|M|^{1/2}}\Big)du \\
&\quad + \int_{u=-\infty}^{u=\frac{X'_{2ik}\beta}{L^{-1/2}}} L^{-1}u\text{exp}\Big\{-\frac{1}{2}u^2\Big\}\Big(Z'_{1ij}\frac{1}{|D|^{1/2}\sqrt{2\pi}}\frac{1}{|M|^{1/2}}(-M^{-1}Z_{2ik})\Big)du
\end{aligned}
$$

Here, we have rewritten the variables

$$\begin{aligned} u &= \frac{s}{L^{-1/2}}, \\ L &= (1-(Z'_{2ik}(D^{-1}+Z_{2ik}Z'_{2ik})^{-1}Z_{2ik})), \\ M &= D^{-1}+\widetilde{Z}_{2ik}\widetilde{Z}'_{2ik}. \end{aligned}$$

Next, integration over u produces

$$E[Y_{1ij}Y_{2ik}] = \frac{1}{|D|^{1/2}}\frac{1}{|M|^{1/2}}\frac{1}{L^{1/2}}X'_{1ij}\beta\Phi(L^{1/2}X'_{2ik}\beta)+\frac{1}{|D|^{1/2}}\frac{1}{|M|^{1/2}}\frac{1}{L}Z'_{1ij}M^{-1}Z_{2ik}\phi(L^{1/2}X'_{2ik}\beta).$$

Hence, the covariance equals

$$\mathrm{Cov}[Y_{1ij}Y_{2ik}] = \left(\frac{1}{|D|^{1/2}}\frac{1}{|M|^{1/2}}\frac{1}{L^{1/2}}-1\right)X'_{1ij}\beta\Phi(L^{1/2}X'_{2ik}\beta)+\frac{1}{|D|^{1/2}}\frac{1}{|M|^{1/2}}\frac{1}{L}Z'_{1ij}M^{-1}Z_{2ik}\phi(L^{1/2}X'_{2ik}\beta).$$

As a result, the correlation is equal to

$$\rho_{Y_{1ij},Y_{2ik}} = \frac{\left(\frac{1}{|D|^{1/2}}\frac{1}{|M|^{1/2}}\frac{1}{L^{1/2}}-1\right)X'_{1ij}\beta\Phi(L^{1/2}X'_{2ik}\beta)+\frac{1}{|D|^{1/2}}\frac{1}{|M|^{1/2}}\frac{1}{L}Z'_{1ij}M^{-1}Z_{2ik}\phi(L^{1/2}X'_{2ik}\beta)}{\sqrt{\left(Z'_{1ij}DZ_{1ij}+\Sigma_{1ij}\right)\Phi(L^{1/2}X'_{2ik}\beta)(1-\Phi(L^{1/2}X'_{2ik}\beta))}}$$

.

G Correlation of vectors of responses

Let us derive the manifest variance-covariance matrix between subvectors of the two responses $\widetilde{Y}_{1i}$ of length $\widetilde{n}_i$ and $\widetilde{Y}_{2i}$ of length $\widetilde{p}_i$. Since the subvectors are independent conditional on the random effects, the expected value of the product equals

$$
\begin{aligned}
&E[Y_{1i}Y_{2i}] \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \Big(\int_{\boldsymbol{t}=-\infty}^{\boldsymbol{t}=X'_{2ik}\beta+\widetilde{Z}_{2i}\boldsymbol{b}_i}\Big)^{\widetilde{p}_i} \frac{(X'_{1ij}\beta+\widetilde{Z}_{1i}\boldsymbol{b}_i)}{(2\pi)^{(q+1)/2}|D|^{1/2}}\exp\Big\{-\tfrac{1}{2}(\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i+\boldsymbol{t}'\boldsymbol{t})\Big\}d\boldsymbol{b}_id\boldsymbol{t} \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \Big(\int_{\boldsymbol{t}=-\infty}^{\boldsymbol{t}=X'_{2ik}\beta+\widetilde{Z}_{2i}\boldsymbol{b}_i}\Big)^{\widetilde{p}_i} \frac{\widetilde{Z}_{1i}\boldsymbol{b}_i}{(2\pi)^{(q+1)/2}|D|^{1/2}}\exp\Big\{-\tfrac{1}{2}(\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i+\boldsymbol{t}'\boldsymbol{t})\Big\}d\boldsymbol{b}_id\boldsymbol{t} \\
&\quad +\widetilde{X}'_{1i}\beta\Big[\Phi(\widetilde{X}^1_{2i}\beta,\tfrac{1}{L^1})\quad \Phi(\widetilde{X}^2_{2i}\beta,\tfrac{1}{L^2})\quad \dots\quad \Phi(\widetilde{X}^{\widetilde{p}_i}_{2i}\beta,\tfrac{1}{L^{\widetilde{p}_i}})\Big]' \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \Big(\int_{\boldsymbol{s}=-\infty}^{\boldsymbol{s}=X'_{2ik}\beta}\Big)^{\widetilde{p}_i} \frac{\widetilde{Z}_{1i}\boldsymbol{b}_i}{(2\pi)^{(q+1)/2}|D|^{1/2}}\exp\Big\{-\tfrac{1}{2}\Big(\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i+\big(\boldsymbol{s}+\widetilde{Z}_{2i}\boldsymbol{b}_i\big)'\big(\boldsymbol{s}+\widetilde{Z}_{2i}\boldsymbol{b}_i\big)\Big)\Big\}d\boldsymbol{b}_id\boldsymbol{s} \\
&\quad +\widetilde{X}'_{1i}\beta\Big[\Phi(\widetilde{X}^1_{2i}\beta,\tfrac{1}{L^1})\quad \Phi(\widetilde{X}^2_{2i}\beta,\tfrac{1}{L^2})\quad \dots\quad \Phi(\widetilde{X}^{\widetilde{p}_i}_{2i}\beta,\tfrac{1}{L^{\widetilde{p}_i}})\Big]' \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \frac{1}{(2\pi)^{(q+1)/2}|D|^{1/2}}\widetilde{Z}_{1i}
\begin{bmatrix}
\int_{s=-\infty}^{s=\widetilde{X}^1_{2i}\beta}\boldsymbol{b}_i\exp\Big\{-\tfrac{1}{2}\Big(\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i+(s+\widetilde{Z}^1_{2i}b_i)^2\Big)\Big\}dsd\boldsymbol{b}_i \\
\int_{s=-\infty}^{s=\widetilde{X}^2_{2i}\beta}\boldsymbol{b}_i\exp\Big\{-\tfrac{1}{2}\Big(\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i+(s+\widetilde{Z}^2_{2i}b_i)^2\Big)\Big\}dsd\boldsymbol{b}_i \\
\dots \\
\int_{s=-\infty}^{s=\widetilde{X}^{\widetilde{p}_i}_{2i}\beta}\boldsymbol{b}_i\exp\Big\{-\tfrac{1}{2}\Big(\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i+(s+\widetilde{Z}^{\widetilde{p}_i}_{2i}b_i)^2\Big)\Big\}dsd\boldsymbol{b}_i
\end{bmatrix}' \\
&\quad +\widetilde{X}'_{1i}\beta\Big[\Phi(\widetilde{X}^1_{2i}\beta,(L^{-1})_1)\quad \Phi(\widetilde{X}^2_{2i}\beta,(L^{-1})_2)\quad \dots\quad \Phi(\widetilde{X}^{\widetilde{p}_i}_{2i}\beta,(L^{-1})_{\widetilde{p}_i})\Big] \\
&= \Big(\int_{-\infty}^{+\infty}\Big)^q \frac{1}{(2\pi)^{(q+1)/2}|D|^{1/2}}\widetilde{Z}_{1i}
\begin{bmatrix}
\int_{s=-\infty}^{s=\widetilde{X}^1_{2i}\beta}\boldsymbol{b}_i\exp\Big\{-\tfrac{1}{2}\Big(\boldsymbol{b}_i-u_1)'E_1^{-1}(\boldsymbol{b}_i-u_1)+O_1\Big)\Big\}dsd\boldsymbol{b}_i \\
\int_{s=-\infty}^{s=\widetilde{X}^2_{2i}\beta}\boldsymbol{b}_i\exp\Big\{-\tfrac{1}{2}\Big(\boldsymbol{b}_i-u_2)'E_2^{-1}(\boldsymbol{b}_i-u_2)+O_2\Big)\Big\}dsd\boldsymbol{b}_i \\
\dots \\
\int_{s=-\infty}^{s=\widetilde{X}^{\widetilde{p}_i}_{2i}\beta}\boldsymbol{b}_i\exp\Big\{-\tfrac{1}{2}\Big(\boldsymbol{b}_i-u_{\widetilde{p}_i})'E_{\widetilde{p}_i}^{-1}(\boldsymbol{b}_i-u_{\widetilde{p}_i})+O_{\widetilde{p}_i}\Big)\Big\}dsd\boldsymbol{b}_i
\end{bmatrix}' \\
&\quad +\widetilde{X}'_{1i}\beta\Big[\Phi(\widetilde{X}^1_{2i}\beta,(L^{-1})_1)\quad \Phi(\widetilde{X}^2_{2i}\beta,(L^{-1})_2)\quad \dots\quad \Phi(\widetilde{X}^{\widetilde{p}_i}_{2i}\beta,(L^{-1})_{\widetilde{p}_i})\Big]
\end{aligned}
$$

where the superscript denotes the row of the submatrix. Here we have rewritten the variables $\boldsymbol{t}=\boldsymbol{s}+\widetilde{Z}_{2i}\boldsymbol{b}_i$ and $\boldsymbol{b}'_iD^{-1}\boldsymbol{b}_i+(s+\widetilde{Z}^2_{2i}b_i)^2=(\boldsymbol{b}_i-u_{\widetilde{p}_i})'E^{-1}_{\widetilde{p}_i}(\boldsymbol{b}_i-u_{\widetilde{p}_i})+O_{\widetilde{p}_i}$, with

$$
\begin{aligned}
E_{\widetilde{p}_i}^{-1} &= \widetilde{Z}^{\widetilde{p}_i}{}'_{2i}\widetilde{Z}_{2i}^{\widetilde{p}_i} + D^{-1},\\
O_{\widetilde{p}_i} &= s^2 - (\widetilde{Z}^{\widetilde{p}_i}{}'_{2i}s)E_{\widetilde{p}_i}(\widetilde{Z}_{2i}^{\widetilde{p}_i}s)',\\
u_{\widetilde{p}_i} &= -E_{\widetilde{p}_i}\widetilde{Z}'_{2i}{}^{\widetilde{p}_i}s.
\end{aligned}
$$

Next, integrating over $\boldsymbol{b}_i$ produces

$$
\begin{aligned}
&E[Y_{1ij}Y_{2ik}]\\
&= \frac{1}{(2\pi)^{1/2}|D|^{1/2}}\widetilde{Z}_{1i}\begin{bmatrix}
\int_{s=-\infty}^{s=\widetilde{X}_{2i}^{1}\beta}|E_1|^{1/2}u_1\mathsf{exp}\Big\{-\frac{1}{2}O_1\Big\}ds\\
\int_{s=-\infty}^{s=\widetilde{X}_{2i}^{2}\beta}|E_2|^{1/2}u_2\mathsf{exp}\Big\{-\frac{1}{2}O_2\Big\}ds\\
\cdots\\
\int_{s=-\infty}^{s=\widetilde{X}_{2i}^{\widetilde{p}_i}\beta}|E_{\widetilde{p}_i}|^{1/2}u_{\widetilde{p}_i}\mathsf{exp}\Big\{-\frac{1}{2}O_{\widetilde{p}_i}\Big\}ds
\end{bmatrix}'\\
&\quad+\widetilde{X}'_{1i}\beta\left[\Phi(\widetilde{X}_{2i}^{1}\beta,\tfrac{1}{L^1})\quad\Phi(\widetilde{X}_{2i}^{2}\beta,\tfrac{1}{L^2})\quad\ldots\quad\Phi(\widetilde{X}_{2i}^{\widetilde{p}_i}\beta,\tfrac{1}{L^{\widetilde{p}_i}})\right]\\
&= \frac{-1}{(2\pi)^{1/2}|D|^{1/2}}\widetilde{Z}_{1i}\begin{bmatrix}
\int_{s=-\infty}^{s=\widetilde{X}_{2i}^{1}\beta}|E_1|^{1/2}E_1\widetilde{Z}^{1}{}'_{2i}s\mathsf{exp}\Big\{-\frac{1}{2}\Big(s(1-\widetilde{Z}_{2i}^{1}E_1\widetilde{Z}^{1}{}'_{2i})s\Big)\Big\}ds\\
\int_{s=-\infty}^{s=\widetilde{X}_{2i}^{2}\beta}|E_2|^{1/2}E_2\widetilde{Z}^{2}{}'_{2i}s\mathsf{exp}\Big\{-\frac{1}{2}\Big(s(1-\widetilde{Z}_{2i}^{2}E_2\widetilde{Z}^{2}{}'_{2i})s\Big)\Big\}ds\\
\cdots\\
\int_{s=-\infty}^{s=\widetilde{X}_{2i}^{\widetilde{p}_i}\beta}|E_{\widetilde{p}_i}|^{1/2}E_{\widetilde{p}_i}\widetilde{Z}^{\widetilde{p}_i}{}'_{2i}s\mathsf{exp}\Big\{-\frac{1}{2}\Big(s(1-\widetilde{Z}_{2i}^{\widetilde{p}_i}E_{\widetilde{p}_i}\widetilde{Z}^{\widetilde{p}_i}{}'_{2i})s\Big)\Big\}ds
\end{bmatrix}'\\
&+\quad\widetilde{X}'_{1i}\beta\left[\Phi(\widetilde{X}_{2i}^{1}\beta,\tfrac{1}{L^1})\quad\Phi(\widetilde{X}_{2i}^{2}\beta,\tfrac{1}{L^2})\quad\ldots\quad\Phi(\widetilde{X}_{2i}^{\widetilde{p}_i}\beta,\tfrac{1}{L^{\widetilde{p}_i}})\right]\\
&= \frac{\widetilde{Z}_{1i}}{\sqrt{|D|}}\begin{bmatrix}
\sqrt{|E_1|}E_1\widetilde{Z}^{1}{}'_{2i}\phi\Big(\widetilde{X}_{2i}^{1}\beta\cdot\sqrt{1-\widetilde{Z}_{2i}^{1}E_1\widetilde{Z}^{1}{}'_{2i}}\Big)\Big(1-\widetilde{Z}_{2i}^{1}E_1\widetilde{Z}^{1}{}'_{2i}\Big)^{-1}\\
\sqrt{|E_2|}E_2\widetilde{Z}^{2}{}'_{2i}\phi\Big(\widetilde{X}_{2i}^{2}\beta\cdot\sqrt{1-\widetilde{Z}_{2i}^{2}E_2\widetilde{Z}^{2}{}'_{2i}}\Big)\Big(1-\widetilde{Z}_{2i}^{2}E_2\widetilde{Z}^{2}{}'_{2i}\Big)^{-1}\\
..\\
\sqrt{|E_{\widetilde{p}_i}|}E_{\widetilde{p}_i}\widetilde{Z}^{\widetilde{p}_i}{}'_{2i}\phi\Big(\widetilde{X}_{2i}^{\widetilde{p}_i}\beta\cdot\sqrt{1-\widetilde{Z}_{2i}^{\widetilde{p}_i}E_{\widetilde{p}_i}\widetilde{Z}^{\widetilde{p}_i}{}'_{2i}}\Big)\Big(1-\widetilde{Z}_{2i}^{p}E_p\widetilde{Z}^{p}{}'_{2i}\Big)^{-1}
\end{bmatrix}'\\
&\quad+\widetilde{X}'_{1i}\beta\left[\Phi(\widetilde{X}_{2i}^{1}\beta,(L^{-1})_1)\quad\Phi\widetilde{X}_{2i}^{2}\beta,(L^{-1})_2)\quad\ldots\quad\Phi(\widetilde{X}_{2i}^{\widetilde{p}_i}\beta,(L^{-1})_{\widetilde{p}_i})\right]'.
\end{aligned}
$$

Hence, the covariance equals

$$\frac{\widetilde{Z}_{1i}}{\sqrt{|D|}}\begin{bmatrix} \sqrt{|E_1|}E_1\widetilde{Z}^{1\prime}{}_{2i}\phi\left(\widetilde{X}^1_{2i}\beta\cdot\sqrt{1-\widetilde{Z}^1_{2i}E_1\widetilde{Z}^{1\prime}{}_{2i}}\right)\left(1-\widetilde{Z}^1_{2i}E_1\widetilde{Z}^{1\prime}{}_{2i}\right)^{-1} \\ \sqrt{|E_2|}E_2\widetilde{Z}^{2\prime}{}_{2i}\phi\left(\widetilde{X}^2_{2i}\beta\cdot\sqrt{1-\widetilde{Z}^2_{2i}E_2\widetilde{Z}^{2\prime}{}_{2i}}\right)\left(1-\widetilde{Z}^2_{2i}E_2\widetilde{Z}^{2\prime}{}_{2i}\right)^{-1} \\ .. \\ \sqrt{|E_{\widetilde{p}_i}|}E_{\widetilde{p}_i}\widetilde{Z}^{\widetilde{p}_i\prime}{}_{2i}\phi\left(\widetilde{X}^{\widetilde{p}_i}_{2i}\beta\cdot\sqrt{1-\widetilde{Z}^{\widetilde{p}_i}_{2i}E_{\widetilde{p}_i}\widetilde{Z}^{\widetilde{p}_i\prime}{}_{2i}}\right)\left(1-\widetilde{Z}^{\widetilde{p}_i}_{2i}E_{\widetilde{p}_i}\widetilde{Z}^{\widetilde{p}_i\prime}{}_{2i}\right)^{-1} \end{bmatrix}'$$