Contents

Supplementary Notes for "Causal effect heterogeneity estimation using summary statistics"

Yadong Yang[1,†], Minxi Bai[1,†], Jiacheng Miao[2], Stephen Dorn[2],

Jonathan Haugstad[2], Jin Liu[3,*], Qiongshi Lu[2,*], Xingjie Shi[1,*]

[1]KLATASDS-MOE, Academy of Statistics and Interdisciplinary Sciences, School of Statistics, East China Normal University

[2]Department of Biostatistics & Medical Informatics, University of Wisconsin-Madison

[3]School of Data Science, The Chinese University of Hong Kong-Shenzhen

January 4, 2026

Contents

*Corresponding author, liujinlab@cuhk.edu.cn, qlu@biostat.wisc.edu, xjshi@fem.ecnu.edu.cn.
† The first two authors have contributed equally to this work.

S1 The MERLIN approach

S1.1 Statistical model for MERLIN

The MERLIN framework employs a linear structural model to define the relationship between an exposure X and an outcome Y. For a given individual i, this model incorporates M genetic variants (G_j, for $j = 1, \ldots, M$ SNPs), an environmental modifier E_i, and an unmeasured common confounder U_i. The individual-level model is specified as: without loss of generality, we denote X_i and Y_i as the covariate-adjusted exposure and outcome for individual i. The individual-level model is specified as:

$$X_i = \sum_j G_{ij}\gamma_j^{(G)} + E_i\gamma^{(E)} + \sum_j G_{ij}E_i\gamma_j^{(GI)} + U_i\eta_X + \epsilon_{X_i}, \tag{1}$$

$$Y_i = X_i\beta^{(A)} + \sum_j G_{ij}\beta_j^{(G)} + E_i\beta^{(E)} + X_iE_i\beta^{(I)} + U_i\eta_Y + \epsilon_{Y_i}. \tag{2}$$

Apart from standard regression assumptions, the validity of MERLIN relies on the following key assumptions.

1. Core assumptions for instrumental variables [1]:
   - Confounder Independence: $G_j \perp\!\!\!\perp U_i$ and $E_i \perp\!\!\!\perp U_i$, $\forall j$.
   - G-E Independence: $G_j \perp\!\!\!\perp E_i$, $\forall j$.
   - Exposure Relevance: $G_j \not\perp\!\!\!\perp X$, and $G_jE \not\perp\!\!\!\perp X$, $\forall j$.
2. Distributional and centering assumptions for model simplification:
   - G_j and E_i are mean-centered ($\mathbb{E}[G_j] = 0$ and $\mathbb{E}[E_i] = 0$).
   - For a discrete binary modifier, we use a standardized form (values $\sqrt{\frac{1-p}{p}}, -\sqrt{\frac{p}{1-p}}$) that ensures $\mathbb{E}[E_i] = 0$ and $\mathbb{E}[E_i]^2 = 1$. Its skewness, $\mu_3 = \mathbb{E}E_i^3$, is generally non-zero unless the binary categories are equally frequent ($p = 0.5$).
   - For a continuous modifier, our simplest derivations assume it has zero skewness ($\mu_3 = \mathbb{E}E_i^3 = 0$), a property of symmetric distributions like the normal distribution.

S1.2 Distribution of summary statistics for the exposure

We aim to establish the approximate distribution for the vector of estimated genetic effects on exposure, $\hat{\boldsymbol{\gamma}} = (\hat{\boldsymbol{\gamma}}^{(G),\top}, \hat{\boldsymbol{\gamma}}^{(GI),\top})^\top$. Here, $\hat{\boldsymbol{\gamma}}_1$ are the estimated main effects of SNPs (G_j) on X, and $\hat{\boldsymbol{\gamma}}^{(GI)}$ are estimated $G_j \times E$ interaction effects on X. These are typically obtained as marginal (single-predictor or interaction-term-specific) ordinary least square (OLS) estimates from GWAS and GWIS regressions, respectively.

Let $W = [G, GE]$ be the $n \times 2M$ matrix of genetic predictors of interest (where G is $n \times M$ and GE is $n \times M$, with its j-th column being $G_{ij}E_i$). Let $\boldsymbol{\gamma} = (\boldsymbol{\gamma}^{(G),\top}, \boldsymbol{\gamma}^{(GI),\top})^\top$ be the vector of true effects of these predictors from the structure model for exposure (Equation 1), which can be written in matrix form as:

$$X = W\boldsymbol{\gamma} + E\gamma^{(E)} + \varepsilon_x, \tag{3}$$

where the residual term is $\varepsilon_x = U\eta_X + \epsilon_X$ from Equation (1).

The vector of marginal OLS estimates $\hat{\boldsymbol{\gamma}}$ is given by $\hat{\boldsymbol{\gamma}} = D_W^{-1}W^\top X$, where $D_W = \text{diag}(W_k^\top W_k)$ is a diagonal matrix with the sum of squares of each predictor W_k (column of W). Substituting Equation (3): $\hat{\boldsymbol{\gamma}} = D_W^{-1}W^\top W\boldsymbol{\gamma} + D_W^{-1}W^\top E\gamma^{(E)} + D_W^{-1}W^\top\varepsilon_x$. We assume that the genetic predictors W are exogenous with respect to the full residual ε_x. Specifically, $\mathbb{E}[W^\top\varepsilon_x] = 0$. This holds if: (i) G_j and E_i are independent of unmeasured confounders U_i, and (ii) G_j and E_i are independent of the errors ϵ_{Xi}. Under these conditions, the expectation of $\hat{\boldsymbol{\gamma}}$ is $\mathbb{E}[\hat{\boldsymbol{\gamma}}] = D_W^{-1}W^\top W\boldsymbol{\gamma} + D_W^{-1}W^\top E\gamma^{(E)}$.

The asymptotic distribution of $\sqrt{n}(\hat{\boldsymbol{\gamma}} - \mathbb{E}[\hat{\boldsymbol{\gamma}}])$ can be derived using the multivariate Lindeberg-Levy Central Limit Theorem for $\frac{1}{\sqrt{n}}W^\top\varepsilon_x$ and Slutsky's Theorem for the product with the scaling term

nD_W^{-1}. The term nD_W^{-1} converges in probability to $B_W = \text{diag}^{-1}(\sigma^2_{W_k})$, where $\sigma^2_{W_k}$ is the population variance of the kth predictor W_k. Assuming $\text{Var}(\varepsilon_x) = \sigma_x^2 \text{I}_n$, this yields:

$$\sqrt{n}(\hat{\boldsymbol{\gamma}} - \mathbb{E}[\hat{\boldsymbol{\gamma}}]) \xrightarrow{d} \mathcal{N}(\mathbf{0}, B_W \Sigma_W B_W \sigma_x^2),$$

where $\Sigma_W = \mathbb{E}[\frac{1}{n} W^\top W]$ is the population covariance matrix of the predictors in W, which can be written as $\Sigma_W = \text{diag}(\sigma_{W_k}) R_W \text{diag}(\sigma_{W_k})$, with R_W being the population correlation matrix of predictors.

Under the stated conditions for E (independent of G and $\mathbb{E}[E_i] = 0$), R_W simplifies to a block diagonal matrix $R_W = \text{diag}(R_{LD}, R_{LD})$, where R_{LD} is the linkage disequilibrium (LD) matrix of G. We note that the block structure of the asymptotical covariance implies that $\hat{\boldsymbol{\gamma}}_1$ and $\hat{\boldsymbol{\gamma}}_3$ are asymptotically independent.

Now we express $\mathbb{E}[\hat{\boldsymbol{\gamma}}]$ and $\text{Var}[\hat{\boldsymbol{\gamma}}]$ using summary statistics. Let $\hat{R}_W = D_W^{-\frac{1}{2}}(W^\top W) D_W^{-\frac{1}{2}}$ be the sample correlation matrix of W. Under the assumption that $G \perp\!\!\!\perp E$, $D_W^{-1} W^\top E \xrightarrow{p} \mathbf{0}$. Then

$$\mathbb{E}[\hat{\boldsymbol{\gamma}}] \approx D_W^{-\frac{1}{2}} D_W^{-\frac{1}{2}} W^\top W D_W^{-\frac{1}{2}} D_W^{\frac{1}{2}} \boldsymbol{\gamma} = D_W^{-\frac{1}{2}} \hat{R}_W D_W^{\frac{1}{2}} \boldsymbol{\gamma}.$$

Let $\boldsymbol{S}_W = \text{diag}(S_1, S_3)$ contains standard errors $se(\hat{\gamma}_{1j})$ and $se(\hat{\gamma}_{3j})$. For marginal regression, $\boldsymbol{S}^2_{W,kk} \approx \sigma_x^2 (D_W^{-1})_{kk}$ (i.e., $\frac{\sigma_x^2}{W_k^\top W_k}$) [2]. This impels $\boldsymbol{S}_W \approx \sigma_x D_W^{-\frac{1}{2}}$. Using these approximations:

$$\mathbb{E}[\hat{\boldsymbol{\gamma}}] \approx (\boldsymbol{S}_W \sigma_x^{-1}) \hat{R}_W (\boldsymbol{S}_W^{-1} \sigma_x) \boldsymbol{\gamma} = \boldsymbol{S}_W \hat{R}_W \boldsymbol{S}_W^{-1} \boldsymbol{\gamma}.$$

As $B_W \Sigma_W B_W^\top = B_W^{\frac{1}{2}} R_W B_W^{\frac{1}{2}}$, it can be approximated by $n^{1/2} D_W^{-1/2} R_W n^{1/2} D_W^{-1/2}$. Using these approximations:

$$\frac{1}{n} B_W \Sigma_W B_W \sigma_x^2 \approx (\boldsymbol{S}_W \sigma_x^{-1}) \hat{R}_W (\boldsymbol{S}_W \sigma_x^{-1}) \sigma_x^2 = \boldsymbol{S}_W \hat{R}_W \boldsymbol{S}_W.$$

In practice, $\hat{R}_W$ is replaced by $\text{diag}(R, R)$, where R is an LD matrix from a suitable reference panel. Let S_1 and S_2 be diagonal matrices of standard errors for $\hat{\boldsymbol{\gamma}}^{(G)}$ and $\hat{\boldsymbol{\gamma}}^{(GI)}$ respectively. The final approximate distributions are:

$$\hat{\boldsymbol{\gamma}}^{(G)} \mid \boldsymbol{\gamma} \overset{A}{\sim} \mathcal{N}(S_1 R S_1^{-1} \boldsymbol{\gamma}^{(G)}, S_1 R S_1),$$
$$\hat{\boldsymbol{\gamma}}^{(GI)} \mid \boldsymbol{\gamma} \overset{A}{\sim} \mathcal{N}(S_2 R S_2^{-1} \boldsymbol{\gamma}^{(GI)}, S_2 R S_2).$$

S1.3 Distribution of summary statistics in the outcome

We now derive the approximate distribution of GWAS and GWIS summary statistics for the outcome, $\hat{\Gamma} = (\hat{\Gamma}^{(G),\top}, \hat{\Gamma}^{(GI),\top})^\top$, where $\hat{\Gamma}^{(G)}$ are the estimated main effects of SNPs (G_j) on outcome Y, and $\hat{\Gamma}^{(GI)}$ are the estimated $G_j \times E$ interaction effects on Y. We assume summary statistics are from cohorts with no overlap with exposure cohorts (see Supplementary Note S2 for overlap adjustments).

Let $Z = [G, GE]$ be the $N \times 2M$ matrix of genetic predictors used in modeling the outcome. It is assumed here that the outcome cohort (from which Z is derived) does not overlap with the exposure cohort (associated with the predictor matrix W), although Z is formed using the same set of M SNPs and the modifier E as used in W. The true underlying composite genetic effects that these summary statistics aim to capture are $\Gamma = (\Gamma^{(G),\top}, \Gamma^{(GI),\top})^\top$, where $\Gamma^{(G)} = \beta^{(A)} \boldsymbol{\gamma}^{(G)} + \boldsymbol{\beta}^{(A)}$, $\Gamma^{(GI)} = \beta^{(A)} \boldsymbol{\gamma}^{(GI)} + \beta^{(I)} \boldsymbol{\gamma}^{(G)}$. By substituting equation (1) into equation (2), the full structural model for the outcome can be written in matrix form as:

$$Y = Z\Gamma + GE^2 \beta^{(I)} \boldsymbol{\gamma}^{(GI)} + E(\beta^{(A)} \gamma^{(E)} + \beta^{(E)}) + E^2 \beta^{(I)} \gamma^{(E)} + \varepsilon_y, \tag{4}$$

where E^2 is an element-wise squaring of the vector E (or a vector of E_i^2), and GE^2 has elements $G_{ij} E_i^2$. The residual term ε_y comprises all terms from the full reduced form not explicitly dependent on G, GE, or GE^2:

$$\varepsilon_y = EU\eta_X \beta^{(I)} + U(\eta_X \beta^{(A)} + \eta_Y) + \epsilon_X \beta^{(A)} + E\epsilon_X \beta^{(I)} + \epsilon_Y.$$

Under conditions for E stated previously, we have $\mathbb{E}[\varepsilon_y] = \mathbf{0}$,$\text{Var}[\varepsilon_y] = \sigma_y^2 \text{I}_N$, and $\mathbb{E}[Z^\top \varepsilon_y] = \mathbf{0}$.

Let $D_Z = \text{diag}(Z_k^T Z_k)$ be the diagonal matrix of sums of squares for columns Z_k of Z. The marginal OLS estimates $\hat{\Gamma} = D_Z^{-1} Z^\top Y$. Thus, the mean

$$\mathbb{E}[\hat{\Gamma}] = D_Z^{-1} Z^\top Z\Gamma + D_Z^{-1} Z^\top G E^2 \beta^{(I)} \boldsymbol{\gamma}^{(GI)} + D_Z^{-1} Z^\top E(\beta^{(A)} \gamma^{(E)} + \beta^{(E)}) + D_Z^{-1} Z^\top E^2 \beta^{(I)} \gamma^{(E)}.$$

The term $\sqrt{N}(\hat{\Gamma} - \mathbb{E}[\hat{\Gamma}]) = (N D_Z^{-1})(\frac{1}{\sqrt{N}} Z^\top \varepsilon_y)$. Following similar arguments as for the exposure statistics (Lindeberg-Levy CLT for $\frac{1}{\sqrt{N}} Z^\top \varepsilon_y$; $N D_Z^{-1} \xrightarrow{p} B_Z = \text{diag}^{-1}(\sigma^2_{Z_k})$; Slutsky's Theorem), we have

$$\sqrt{N}(\hat{\Gamma} - \mathbb{E}[\hat{\Gamma}]) \xrightarrow{d} \mathcal{N}(\mathbf{0}, B_Z \Sigma_Z B_Z \sigma_y^2),$$

where $\Sigma_Z = \mathbb{E}[\frac{1}{N} Z^\top Z]$ is the population covariance matrix of predictors in Z. As $G_j \perp\!\!\!\perp E_i$, then $\Sigma_Z = \text{diag}(\Sigma_{Z1}, \Sigma_{Z2})$. This implies

$$\frac{1}{N} B_Z \Sigma_Z B_Z \sigma_y^2 \approx \boldsymbol{S}_Z R_Z \boldsymbol{S}_Z,$$

where $\boldsymbol{S}_Z = \text{diag}(S_3, S_4)$ contains standard errors $se(\hat{\Gamma}_j^{(G)})$ and $se(\hat{\Gamma}_j^{(GI)})$, and R_Z is correlation matrix of Z (approximated by $\text{diag}(R_{LD}, R_{LD})$). We note that, based on the block structure of the asymptotical covariance, GWAS $\hat{\Gamma}^{(G)}$ and GWIS $\hat{\Gamma}^{(GI)}$ for the outcome are asymptotically independent.

The crucial step is to define $\mathbb{E}[\hat{\Gamma}]$ using summary statistics. Given $\mathbb{E}[G_j] = 0$ and $\mathbb{E}[E_i] = 0$, term like $D_Z^{-1} Z^\top E$ and $D_Z^{-1} Z^\top E^2$ will converge to zero with probability one, simplifying

$$\mathbb{E}[\hat{\Gamma}] \approx D_Z^{-1} Z^\top Z\Gamma + D_Z^{-1} Z^\top G E^2 \beta^{(I)} \boldsymbol{\gamma}^{(GI)}.$$

The first term $D_Z^{-1} Z^\top Z\Gamma \approx \boldsymbol{S}_Z R_Z \boldsymbol{S}_Z^{-1} \Gamma$. The second term $D_Z^{-1} Z^\top G E^2 \beta^{(I)} \boldsymbol{\gamma}^{(GI)}$ can be written in block form as $\begin{bmatrix} D_{Z1}^{-1} G^\top G E^2 \\ D_{Z2}^{-1} G E^\top G E^2 \end{bmatrix} \beta^{(I)} \boldsymbol{\gamma}^{(GI)}$. The lower block $D_{Z2}^{-1} G E^\top G E^2$ involves $\mathbb{E}[E_i^3] = \mu_3$, and can be approximated as

$$\begin{aligned} D_{Z2}^{-1} G E^\top G E^2 &\approx N D_{Z2}^{-1} \Sigma_{Z2} \mathbb{E}[E_i^3] \\ &= N D_{Z2}^{-1} B_{Z1}^{1/2} B_{Z1}^{-1/2} \Sigma_{Z2} B_{Z1}^{-1/2} B_{Z1}^{1/2} \mu_3 \\ &\approx S_4^2 S_3^{-1} R S_3^{-1} \mu_3. \end{aligned}$$

Approximations of these blocks depend on the nature of E.

S1.3.1 Discrete binary environment

For the discrete standardized Bernoulli modifier E_i (values $\sqrt{\frac{1-p}{p}}, -\sqrt{\frac{p}{1-p}}$), we have $\mathbb{E}[E_i] = 0$, $\mathbb{E}[E_i^2] = 1$ and $\mu_3 = \frac{1-2p}{\sqrt{p(1-p)}}$.

Since $\mathbb{E}[E_i^2] = 1$, $\frac{1}{N} G^\top G E^2 \xrightarrow{p} \Sigma_{Z1} \mathbb{E}[E_i^2] = \Sigma_{Z1}$. So the upper block $D_{Z1}^{-1} G^\top G E^2$ can be approximated using summary statistics and the LD reference matrix R as $S_3 R S_3^{-1}$.

Consequently, for large N, the mean for $\hat{\Gamma}$ when E is binary is approximately:

$$\mathbb{E}\hat{\Gamma} \approx \begin{bmatrix} S_3 R S_3^{-1} \Gamma^{(G)} \\ S_4 R S_4^{-1} \Gamma^{(GI)} \end{bmatrix} + \beta^{(I)} \begin{bmatrix} S_3 R S_3^{-1} \\ \mu_3 S_4^2 S_3^{-1} R S_3^{-1} \end{bmatrix} \boldsymbol{\gamma}^{(GI)}.$$

Thus, the approximate distribution are

$$\begin{aligned} \hat{\Gamma}^{(G)} \mid \boldsymbol{\gamma} &\overset{A}{\sim} \mathcal{N}(S_3 R S_3^{-1} (\Gamma^{(G)} + \beta^{(I)} \boldsymbol{\gamma}^{(GI)}), S_3 R S_3), \\ \hat{\Gamma}^{(GI)} \mid \boldsymbol{\gamma} &\overset{A}{\sim} \mathcal{N}\left(S_4 R S_4^{-1} \Gamma^{(GI)} + \mu_3 \beta^{(I)} S_4^2 S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}^{(GI)}, S_4 R S_4 \right). \end{aligned} \tag{5}$$

If $p = 0.5$ (e.g., the male and female have similar sample size), then $E_i \in \{\pm 1\}$, $\mu_3 = 0$, and the distribution of $\hat{\Gamma}$ simplifies:

$$\begin{aligned} \hat{\Gamma}^{(G)} \mid \boldsymbol{\gamma} &\overset{A}{\sim} \mathcal{N}(S_3 R S_3^{-1} (\Gamma^{(G)} + \beta^{(I)} \boldsymbol{\gamma}^{(GI)}), S_3 R S_3), \\ \hat{\Gamma}^{(GI)} \mid \boldsymbol{\gamma} &\overset{A}{\sim} \mathcal{N}(S_4 R S_4^{-1} \Gamma^{(GI)}, S_4 R S_4). \end{aligned}$$

S1.3.2 Continuous environment

Note that the sample correlation matrix of GE is $D_{Z_2}^{-\frac{1}{2}} G^\top GE^2 D_{Z_2}^{-\frac{1}{2}} = D_{Z_2}^{-\frac{1}{2}} GE^\top GE D_{Z_2}^{-\frac{1}{2}} \approx R$ and $\boldsymbol{S}_Z \approx \sigma_y D_Z^{-\frac{1}{2}}$. Based on these approximation, we can estimate $D_Z^{-1} Z^\top GE^2$ using summary statistics as follows:

$$\begin{bmatrix} D_{Z1}^{-1} G^\top GE^2 \\ D_{Z2}^{-1} GE^\top GE^2 \end{bmatrix} \approx \begin{bmatrix} D_{Z1}^{-1} D_{Z2}^{\frac{1}{2}} R D_{Z2}^{\frac{1}{2}} \\ \mu_3 S_4^2 S_3^{-1} R S_3^{-1} \end{bmatrix} \approx \begin{bmatrix} S_3^2 S_4^{-1} R S_4^{-1} \\ \mu_3 S_4^2 S_3^{-1} R S_3^{-1} \end{bmatrix}$$

With this derivation, the approximate distributions are:

$$\begin{aligned} \hat{\Gamma}^{(G)} \mid \boldsymbol{\gamma} &\overset{A}{\sim} \mathcal{N}(S_3 R S_3^{-1} \Gamma^{(G)} + S_3^2 S_4^{-1} R S_4^{-1} \beta^{(I)} \boldsymbol{\gamma}^{(GI)}, S_3 R S_3), \\ \hat{\Gamma}^{(GI)} \mid \boldsymbol{\gamma} &\overset{A}{\sim} \mathcal{N}(S_4 R S_4^{-1} \Gamma^{(GI)} + \mu_3 \beta^{(I)} S_4^2 S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}^{(GI)}, S_4 R S_4). \end{aligned}$$

For continuous E, we assume $\mu_3 = 0$, then the approximate distributions simplifies:

$$\begin{aligned} \hat{\Gamma}^{(G)} \mid \boldsymbol{\gamma} &\overset{A}{\sim} \mathcal{N}(S_3 R S_3^{-1} \Gamma^{(G)} + S_3^2 S_4^{-1} R S_4^{-1} \beta^{(I)} \boldsymbol{\gamma}^{(GI)}, S_3 R S_3), \\ \hat{\Gamma}^{(GI)} \mid \boldsymbol{\gamma} &\overset{A}{\sim} \mathcal{N}(S_4 R S_4^{-1} \Gamma^{(GI)}, S_4 R S_4). \end{aligned} \tag{6}$$

S1.4 Assumption on the Skewness of the Modifier E

In our derivation for the distribution of the outcome GWIS summary statistics ($\hat{\Gamma}^{(GI)}$), we showed that the mean, $\mathbb{E}[\hat{\Gamma}^{(GI)}]$, contains a term dependent on the third moment (skewness) of the modifier E:

$$\mathbb{E}[\hat{\Gamma}^{(GI)}] \approx S_4 R S_4^{-1} \Gamma^{(GI)} + \mu_3 \beta^{(I)} S_4^2 S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}^{(GI)}.$$

For simplicity in our primary analyses and simulations, particularly when estimating the magnitude of non-zero effects and assessing statistical power, we assumed that E has zero skewness ($\mu_3 = 0$), as is the case for any symmetrically distributed modifier (e.g., a standardized normal distribution for continuous E, or a balanced binary distribution where $p = 0.5$).

However, it is critical to note that the term involving skewness is a product: $\mu_3 \times \beta^{(I)}$. Therefore, under the null hypothesis of no interaction effect ($H_0 : \beta^{(I)} = 0$), this entire term vanishes regardless of the value of μ_3. Consequently, the MERLIN test for the presence of an interaction effect ($\beta^{(I)} = 0$) is expected to maintain a correctly calibrated Type I error rate (i.e., not generate excess false positives) even if the modifier E is skewed.

While the significance test for $\beta^{(I)} = 0$ is robust to skewness in E, obtaining an unbiased point estimate of $\beta^{(I)}$ when it is truly non-zero does require accounting for μ_3. For the analyses presented in this paper involving a continuous modifier (age), we proceeded under the simplifying and common assumption that the modifier's distribution is approximately symmetric ($\mu_3 \approx 0$). The MERLIN framework could be extended to incorporate a known non-zero μ_3 for unbiased point estimation in settings with skewed modifiers. For discrete binary modifiers where the categories are not balanced ($p \neq 0.5$), our full likelihood derivation (Eq. (5)) explicitly accounts for the non-zero μ_3.

S2 Addressing sample overlap in summary statistics

In practice, summary statistics from Genome-Wide Association Studies (GWAS) and Genome-Wide Interaction Studies (GWIS) may be derived from distinct cohorts that have partially overlapping participants. This section details the derivation of covariance terms between different sets of summary statistics due to such sample overlap. We assume the standard conditions for the environmental modifier E (independent of genotypes G, and mean-centered $\mathbb{E}[E_i] = 0$) hold, which simplifies several cross-product expectations involving E.

S2.1 Covariance between Exposure GWAS and Exposure GWIS Statistics from Partially Overlapping Cohorts

Let $\hat{\boldsymbol{\gamma}}_{(1)}^{(G)}$ be the exposure GWAS estimates (main G effects on X) from cohort 1 (size N_1) and $\hat{\boldsymbol{\gamma}}_{(2)}^{(GI)}$ be the exposure GWIS estimates (G×E effects on X) from cohort 2 (size N_2), with d_{12} overlapping

individuals. The relevant residuals for exposure X, after accounting for systematic genetic effects ($G\boldsymbol{\gamma}^{(G)}, GE\boldsymbol{\gamma}^{(GI)}$) and the main effect of E ($E\gamma^{(E)}$), are $\varepsilon_{x,(c)} = U_{(c)}\eta_X + \epsilon_{X,(c)}$ for cohort c.

From the general form $\hat{\boldsymbol{\gamma}} - \mathbb{E}[\hat{\boldsymbol{\gamma}}] = D_W^{-1} W^\top \varepsilon_x$, the terms contributing to the covariance are:

$$\hat{\boldsymbol{\gamma}}^{(G)}_{(1)} - \mathbb{E}[\hat{\boldsymbol{\gamma}}^{(G)}_{(1)}] = D^{-1}_{G_{(1)}} {G_{(1)}}^\top \varepsilon_{x,(1)},$$
$$\hat{\boldsymbol{\gamma}}^{(GI)}_{(2)} - \mathbb{E}[\hat{\boldsymbol{\gamma}}^{(GI)}_{(2)}] = D^{-1}_{GE_{(2)}} {GE_{(2)}}^\top \varepsilon_{x,(2)}.$$

where $D_{G_{(1)}} = \mathrm{diag}((G_{(1)})^\top G_{(1)})$ and $D_{GE_{(2)}} = \mathrm{diag}((GE_{(2)})^\top GE_{(2)})$. The residuals $\varepsilon_{x,(1)}$ and $\varepsilon_{x,(3)}$ are from the true underlying model for cohort 1 and 2 respectively (i.e. $\varepsilon_{x,(c)} = U_{(c)}\eta_X + \epsilon_{X,(c)}$).

The covariance between $\hat{\boldsymbol{\gamma}}^{(G)}_{(1)}$ and $\hat{\boldsymbol{\gamma}}^{(GI)}_{(2)}$ (conditional on true effects $\boldsymbol{\gamma}$, genotypes G, and modifier E) is:

$$\mathrm{Cov}(\hat{\boldsymbol{\gamma}}^{(G)}_{(1)}, \hat{\boldsymbol{\gamma}}^{(GI)}_{(2)}) = D^{-1}_{G_{(1)}} {G_{(1)}}^\top \mathrm{Cov}(\varepsilon_{x,(1)}, \varepsilon_{x,(2)}) GE_{(2)} D^{-1}_{G_{(2)}}.$$

The $N_1 \times N_2$ matrix $\mathrm{Cov}(\varepsilon_{x,(1)}, \varepsilon_{x,(2)})$ describes the covariance of residuals between individuals in cohort 1 and cohort 2. If an individual i from cohort 1 is the same as individual i' from cohort 2 (i.e., they are in the overlap of size d_{12}), then $\mathrm{Cov}(\varepsilon_{x,(1)}, \varepsilon_{x,(2)}) = \sigma_x^2$. If individuals i and i' are different, this covariance is 0 (assuming independence of residuals across distinct individuals). Thus, $\mathrm{Cov}(\varepsilon_{x,(1)}, \varepsilon_{x,(2)})$ is a matrix that is $\sigma_x^2 \mathrm{I}_{d_{12}}$. Therefore, we have

$$\begin{aligned}
\mathrm{Cov}(\hat{\boldsymbol{\gamma}}^{(G)}_{(1)}, \hat{\boldsymbol{\gamma}}^{(GI)}_{(2)}) &= \sigma_x^2 D^{-1}_{G_{(1)}} {G_{(o)}}^\top GE_{(o)} D^{-1}_{G_{(2)}} \\
&= \sigma_x D^{-1/2}_{G_{(1)}} D^{-1/2}_{G_{(1)}} D^{1/2}_{G_{(o)}} D^{-1/2}_{G_{(o)}} {G_{(o)}}^\top GE_{(o)} D^{-1/2}_{G_{(o)}} D^{1/2}_{G_{(o)}} D^{-1/2}_{G_{(2)}} D^{-1/2}_{G_{(2)}} \sigma_x
\end{aligned}$$

For large N_1, N_2, and d_{12}, we have $D^{-1}_{G_{(1)}} D_{G_{(o)}} = \mathrm{diag}\left(\frac{\frac{d_{12}}{d_{12}} G^\top_{(0),k} G_{(0),k}}{\frac{N_1}{N_1} G^\top_{(1),k} G_{(1),k}}\right) \xrightarrow{p} \mathrm{diag}\left(\frac{d_{12}\sigma^2_{W_k}}{N_1 \sigma^2_{W_k}} = \frac{d_{12}}{N_1}\right)$, and $D^{-1/2}_{G_{(o)}} {G_{(o)}}^\top GE_{(o)} D^{-1/2}_{G_{(o)}} \xrightarrow{p} \mathbf{0}$.

Consequently, the covariance term is approximately zero:

$$\mathrm{Cov}(\hat{\boldsymbol{\gamma}}^{(G)}_{(1)}, \hat{\boldsymbol{\gamma}}^{(GI)}_{(2)}) = \mathbf{0}.$$

This indicates that exposure GWAS and exposure GWIS summary statistics can be treated as approximately uncorrelated, even if derived from partially overlapping cohorts, provided $G \perp\!\!\!\perp E$ and E is mean-centered. If derived from completely nonoverlapping cohorts ($d_{12} = 0$), this covariance is exactly zero.

S2.2 Covariance between Exposure Statistics and Outcome Statistics with Overlap

Let $\hat{\boldsymbol{\gamma}}_{(exp)}$ denote an exposure vector (either $\hat{\boldsymbol{\gamma}}^{(G)}_{(A)}$ from cohort A using predictors $W_A = G_{(A)}$, or $\hat{\boldsymbol{\gamma}}^{(GI)}_{(A)}$ from cohort A using predictors $W_A = GE_{(A)}$.

Let $\hat{\Gamma}_{(out)}$ denote an outcome vector (either $\hat{\Gamma}^{(G)}_{(B)}$ from cohort B using predictors $W_B = G_{(B)}$, or $\hat{\Gamma}^{(GI)}_{(B)}$ from cohort B using predictors $W_B = GE_{(B)}$.

The general covariance form due to d_{AB} overlapping individuals is:

$$\mathrm{Cov}(\hat{\boldsymbol{\gamma}}_{(exp)}, \hat{\Gamma}_{(out)}) = D^{-1}_{WA} {W_{A,(o)}}^\top \mathrm{Cov}(\varepsilon_{x,(o)}, \varepsilon_{y,(o)}) W_{B,(o)} D^{-1}_{WB},$$

where $\mathrm{Cov}(\varepsilon_{x,(o)}, \varepsilon_{y,(o)}) = \sigma_x \sigma_y \rho_{xy}$ is generally non-zero due to shared unmeasured confounders U_i and the propagation of exposure-specific errors ϵ_{Xi} into $\varepsilon_{x,(o)}$.

1. Exposure GWAS vs. Outcome GWAS: $W_A = G_{(1)}$, $W_B = G_{(3)}$,

$$\begin{aligned}
\mathrm{Cov}(\hat{\boldsymbol{\gamma}}^{(G)}_{(1)}, \hat{\Gamma}^{(G)}_{(3)}) &= \sigma_x \sigma_y \rho_{xy} D^{-1}_{G_{(1)}} \left[(G_{(o)})^\top G_{(o)}\right] D^{-1}_{G_{(3)}} \\
&= \rho_{xy} \sigma_x D^{-1/2}_{G_{(1)}} D^{-1/2}_{G_{(1)}} D^{1/2}_{G_{(o)}} D^{-1/2}_{G_{(o)}} {G_{(o)}}^\top G_{(o)} D^{-1/2}_{G_{(o)}} D^{1/2}_{G_{(o)}} D^{-1/2}_{G_{(3)}} D^{-1/2}_{G_{(3)}} \sigma_y \\
&\approx \frac{\rho_{xy} d_{13}}{\sqrt{N_1 N_3}} \boldsymbol{S}_1 R \boldsymbol{S}_3.
\end{aligned}$$

2. Exposure GWAS vs. Outcome GWIS: $W_A = G_{(1)}$, $W_B = GE_{(4)}$, $\text{Cov}(\hat{\boldsymbol{\gamma}}_{(1)}^{(G)}, \hat{\Gamma}_{(4)}^{(GI)}) \approx \mathbf{0}$.

3. Exposure GWIS vs. Outcome GWAS: $W_A = GE_{(2)}$, $W_B = G_{(3)}$, $\text{Cov}(\hat{\boldsymbol{\gamma}}_{(2)}^{(GI)}, \hat{\Gamma}_{(3)}^{(G)}) \approx \mathbf{0}$.

4. Exposure GWIS vs. Outcome GWIS: $W_A = GE_{(2)}$, $W_B = GE_{(4)}$,

$$\begin{aligned}\text{Cov}(\hat{\boldsymbol{\gamma}}_{(2)}^{(GI)}, \hat{\Gamma}_{(4)}^{(GI)}) &= \sigma_x \sigma_y \rho_{xy} D_{GE_{(2)}}^{-1} \left[(GE_{(o)})^\top GE_{(o)}\right] D_{GE_{(4)}}^{-1} \\ &= \rho_{xy}\sigma_x D_{GE_{(2)}}^{-1/2} D_{GE_{(2)}}^{-1/2} D_{GE_{(o)}}^{1/2} D_{GE_{(o)}}^{-1/2} GE_{(o)}{}^\top GE_{(o)} D_{GE_{(o)}}^{-1/2} D_{GE_{(o)}}^{1/2} D_{GE_{(o)}}^{-1/2} D_{GE_{(4)}}^{-1/2} D_{GE_{(4)}}^{-1/2} \sigma_y \\ &\approx \frac{\rho_{xy} d_{24}}{\sqrt{N_2 N_4}} \boldsymbol{S}_2 R \boldsymbol{S}_4.\end{aligned}$$

S2.3 Covariance between Outcome GWAS and Outcome GWIS Statistics from Partially Overlapping Cohorts

$W_A = G_{(3)}$, $W_B = GE_{(4)}$, as shown previously,

$$\text{Cov}(\hat{\boldsymbol{\gamma}}_{(2)}^{(GI)}, \hat{\Gamma}_{(4)}^{(GI)}) = \sigma_y^2 D_{G_{(3)}}^{-1} \left[(G_{(o)})^\top GE_{(o)}\right] D_{GE_{(4)}}^{-1} \approx \mathbf{0}.$$

S2.4 Likelihood Function for MERLIN

Based on the derivations outlined in previous sections, we construct the approximate joint variance-covariance matrix for the full set of four input summary statistic vectors $\hat{\boldsymbol{V}} = (\hat{\boldsymbol{\gamma}}^\top, \hat{\Gamma}^\top)^\top$. Its approximate variance-covariance matrix is a key component of the MERLIN likelihood function and is given by:

$$\text{Cov}[\boldsymbol{V}] \approx \begin{bmatrix} S_1 R S_1 & \mathbf{0} & \rho_1 S_1 R S_3 & \mathbf{0} \\ \mathbf{0} & S_2 R S_2 & \mathbf{0} & \rho_2 S_2 R S_4 \\ \rho_1 S_1 R S_3 & \mathbf{0} & S_3 R S_3 & \mathbf{0} \\ \mathbf{0} & \rho_2 S_2 R S_4 & \mathbf{0} & S_4 R S_4 \end{bmatrix}.$$

where $\rho_1 = \frac{\rho_{xy} d_{13}}{\sqrt{N_1 N_3}}$ and $\rho_2 = \frac{\rho_{xy} d_{24}}{\sqrt{N_2 N_4}}$. The two scalars ρ_1 and ρ_2 quantify the impact of sample overlap, and can be estimated using summary statistics among independent variants with no associations to both exposure and outcome. R is the LD correlation matrix among the M SNPs, typically estimated from a reference panel.

The MERLIN likelihood function assumes that the vector of observed summary statistics $\boldsymbol{V}$ follows a multivariate normal distribution, conditional on the true underlying genetic effects and causal parameters:

$$\hat{\boldsymbol{V}} \overset{A}{\sim} \mathcal{N}(\mathbb{E}[\hat{\boldsymbol{V}}], \text{Cov}[\hat{\boldsymbol{V}}]), \tag{7}$$

where $\mathbb{E}[\hat{\boldsymbol{V}}]$ is the vector of expected values for the summary statistics derived in previous sections, which are functions of $\boldsymbol{\gamma}$, Γ, and thus ultimately of the model parameters $\beta^{(A)}$, $\boldsymbol{\beta}^{(G)}$ and $\beta^{(I)}$. This likelihood is utilized within the Bayesian hierarchical model when all four sets of summary statistics $\mathcal{D} = \{\hat{\boldsymbol{\gamma}}^{(G)}, S_1; \hat{\boldsymbol{\gamma}}^{(GI)}, S_2; \hat{\Gamma}^{(G)}, S_3; \hat{\Gamma}^{(GI)}, S_4\}$ are available and sample overlap between exposure and outcome studies is accounted for through ρ_1 and ρ_2.

S2.5 Estimation of correlation parameters accounting for sample overlap

The parameters ρ_1 and ρ_2 in MERLIN likelihood's covariance matrix quantify the correlation between summary statistics arising from sample overlap. These are estimated empirically using summary statistics from a set of independent "null" genetic variants, i.e., SNPs not associated with either the exposure or the outcome traits involved in the specific covariance term.

To estimate ρ_1, we consider the joint distribution of z-scores for SNP j using the following truncated bivariate normal distribution [3]:

$$\begin{pmatrix} \hat{\boldsymbol{\gamma}}_j^{(G)}/s_{1j} \\ \hat{\Gamma}_j^{(G)}/s_{3j} \end{pmatrix} \mid (\gamma_j^{(G)} = 0, \Gamma_j^{(G)} = 0) \overset{A}{\sim} \mathcal{TN}(\boldsymbol{\mu}, \Sigma, \boldsymbol{a}, \boldsymbol{b}),$$

where $\boldsymbol{\mu} = (0,0)^\top$, and covariance matrix

$$\Sigma = \begin{bmatrix} \varsigma_1^2 & \rho_1 \varsigma_1 \varsigma_2 \\ \rho_1 \varsigma_1 \varsigma_2 & \varsigma_2^2 \end{bmatrix}.$$

The truncation points as well as lower and upper truncation points are fixed at $\boldsymbol{a} = (-1.96, -1.96)^\top$, and $\boldsymbol{b} = (1.96, 1.96)^\top$ are used to select SNPs that are not significantly associated with neither the exposure nor the outcome at a nominal p-value threshold (e.g., $P = 0.05$). This parameter ρ_1 is estimated from a set of LD-pruned "null" SNPs using a Gibbs sampler suitable for truncated multivariate normal distributions [4]. A similar approach is employed to estimate ρ_2.

S3 Parameter Estimation and Statistical Inference for MERLIN

We model the joint log-likelihood for the observed summary statistics, $\mathcal{D}$, using the approximate multivariate normal distribution described in the preceding sections. To simplify notation and improve readability, we set $\boldsymbol{\gamma}^{(G)} = \boldsymbol{\gamma}_1, \boldsymbol{\gamma}^{(GI)} = \boldsymbol{\gamma}_3, \Gamma^{(G)} = \Gamma_1, \Gamma^{(GI)} = \Gamma_3, \beta^{(A)} = \beta_1, \beta^{(I)} = \beta_4, \boldsymbol{\beta}^{(G)} = \boldsymbol{\beta}_2$. For a MERLIN model with a balanced binary E and sample overlap, the full log-likelihood function, which incorporates the assigned prior distributions for the true per-SNP genetic effects, is given by:

$$\begin{aligned} L(\theta|\mathcal{D}) = & \log \mathcal{N}\left(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_1 \\ \hat{\Gamma}_1 \end{pmatrix}; \begin{pmatrix} S_1 R S_1^{-1} \boldsymbol{\gamma}_1 \\ S_3 R S_3^{-1} (\Gamma_1 + \beta_4 \boldsymbol{\gamma}_3) \end{pmatrix}, \begin{pmatrix} S_1 R S_1 & \rho_1 S_1 R S_3 \\ \rho_1 S_1 R S_3 & S_3 R S_3 \end{pmatrix}\right) \\ & + \log \mathcal{N}\left(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_3 \\ \hat{\Gamma}_3 \end{pmatrix}; \begin{pmatrix} S_2 R S_2^{-1} \boldsymbol{\gamma}_3 \\ S_4 R S_4^{-1} \Gamma_3 \end{pmatrix}, \begin{pmatrix} S_2 R S_2 & \rho_2 S_2 R S_4 \\ \rho_2 S_2 R S_4 & S_4 R S_4 \end{pmatrix}\right) \\ & + \log \mathcal{N}(\boldsymbol{\gamma}_1; \mathbf{0}, \sigma_1^2 \mathrm{I}) + \log \mathcal{N}(\boldsymbol{\beta}_2; 0, \sigma_2^2 \mathrm{I}) + \log \mathcal{N}(\boldsymbol{\gamma}_3; \mathbf{0}, \sigma_3^2 \mathrm{I}). \end{aligned} \tag{8}$$

We estimate the model parameters $\theta = (\beta_1, \beta_4)$ using an efficient Gibbs sampling algorithm. This iterative method generates samples whose distribution converges to the target distribution as the sample size increases.

S3.1 Gibbs sampler for MERLIN with sample overlaps

We present the details on deriving Gibbs sampler for MERLIN.

The full conditional distribution of $\boldsymbol{\gamma}$ can be written as

$$\begin{aligned} & \log \Pr(\boldsymbol{\gamma} | \hat{\boldsymbol{\gamma}}, \hat{\Gamma}, \beta_1, \boldsymbol{\beta}_2, \beta_4, \sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_x^2, \sigma_y^2) \\ \propto & \log \Pr(\hat{\Gamma}, \hat{\boldsymbol{\gamma}} \mid \boldsymbol{\gamma}, \beta_1, \boldsymbol{\beta}_2, \beta_4, \sigma_x^2, \sigma_y^2) + \log \Pr(\boldsymbol{\gamma} \mid \sigma_1^2, \sigma_3^2) \\ = & \log \mathcal{N}\left(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_1 \\ \hat{\Gamma}_1 \end{pmatrix}; \begin{pmatrix} S_1 R S_1^{-1} \boldsymbol{\gamma}_1 \\ S_3 R S_3^{-1} (\Gamma_1 + \beta_4 \boldsymbol{\gamma}_3) \end{pmatrix}, \begin{pmatrix} S_1 R S_1 & \rho_1 S_1 R S_3 \\ \rho_1 S_1 R S_3 & S_3 R S_3 \end{pmatrix}\right) \\ & + \log \mathcal{N}\left(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_3 \\ \hat{\Gamma}_3 \end{pmatrix}; \begin{pmatrix} S_2 R S_2^{-1} \boldsymbol{\gamma}_3 \\ S_4 R S_4^{-1} \Gamma_3 \end{pmatrix}, \begin{pmatrix} S_2 R S_2 & \rho_2 S_2 R S_4 \\ \rho_2 S_2 R S_4 & S_4 R S_4 \end{pmatrix}\right) \\ & + \log \mathcal{N}(\boldsymbol{\gamma}_1; \mathbf{0}, \sigma_1^2 \mathrm{I}) + \log \mathcal{N}(\boldsymbol{\gamma}_3; \mathbf{0}, \sigma_3^2 \mathrm{I}) \end{aligned}$$

It follows that the full conditional distribution of γ_{1j} is

$$
\begin{aligned}
&\log\mathcal{N}(\begin{pmatrix}\hat{\boldsymbol{\gamma}}_1\\ \hat{\Gamma}_1\end{pmatrix};\begin{pmatrix}S_1RS_1^{-1}\boldsymbol{\gamma}_1\\ S_3RS_3^{-1}(\Gamma_1+\beta_4\boldsymbol{\gamma}_3)\end{pmatrix},\begin{pmatrix}S_1RS_1 & \rho_1S_1RS_3\\ \rho_1S_1RS_3 & S_3RS_3\end{pmatrix})\\
&+\log\mathcal{N}(\begin{pmatrix}\hat{\boldsymbol{\gamma}}_3\\ \hat{\Gamma}_3\end{pmatrix};\begin{pmatrix}S_2RS_2^{-1}\boldsymbol{\gamma}_3\\ S_4RS_4^{-1}\Gamma_3\end{pmatrix},\begin{pmatrix}S_2RS_2 & \rho_2S_2RS_4\\ \rho_2S_2RS_4 & S_4RS_4\end{pmatrix})\\
&+\log\mathcal{N}(\boldsymbol{\gamma}_1;\mathbf{0},\sigma_1^2\mathrm{I})+\log\mathcal{N}(\boldsymbol{\gamma}_3;\mathbf{0},\sigma_3^2\mathrm{I})\\
\propto&-\frac{1}{2(1-\rho_1^2)}[\boldsymbol{\gamma}_1^\top S_1^{-1}RS_1^{-1}\boldsymbol{\gamma}_1-2\boldsymbol{\gamma}_1^\top S_1^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1+2\rho_1(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1+2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\boldsymbol{\gamma}_1]\\
&-\frac{1}{2(1-\rho_1^2)}[-2\rho_1(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1-2(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}S_3^{-1}\hat{\Gamma}_1]\\
&-\frac{1}{2(1-\rho_1^2)}[(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}RS_3^{-1}(\Gamma_1+\beta_4\boldsymbol{\gamma}_3)]\\
&-\frac{1}{2(1-\rho_2^2)}(-2\boldsymbol{\gamma}_3^\top S_2^{-1}S_2^{-1}\hat{\boldsymbol{\gamma}}_3+2\rho_2\Gamma_3^\top S_2^{-1}S_4^{-1}\hat{\boldsymbol{\gamma}}_3+\boldsymbol{\gamma}_3^\top S_2^{-1}RS_2^{-1}\boldsymbol{\gamma}_3)\\
&-\frac{1}{2(1-\rho_2^2)}(2\rho_2\hat{\Gamma}_3^\top S_2^{-1}S_4^{-1}\boldsymbol{\gamma}_3-2\rho_2\Gamma_3^\top S_2^{-1}RS_4^{-1}\boldsymbol{\gamma}_3-2\Gamma_3^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3+\Gamma_3^\top S_4^{-1}RS_4^{-1}\Gamma_3)\\
&-\frac{\gamma_{1j}^2}{2\sigma_1^2}\\
\propto&-\frac{1}{2(1-\rho_1^2)}(\boldsymbol{\gamma}_1^\top S_1^{-1}RS_1^{-1}\boldsymbol{\gamma}_1-2\rho_1\beta_1\boldsymbol{\gamma}_1^\top S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1+\beta_1^2\boldsymbol{\gamma}_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_1)\\
&-\frac{1}{2(1-\rho_1^2)}(2\rho_1\beta_1\boldsymbol{\gamma}_1^\top S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1-2\beta_1\boldsymbol{\gamma}_1^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1-2\boldsymbol{\gamma}_1^\top S_1^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1+2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\boldsymbol{\gamma}_1)\\
&-\frac{1}{2(1-\rho_1^2)}(-2\rho_1\boldsymbol{\beta}_2^\top S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1-2\rho_1\beta_4\boldsymbol{\gamma}_3^\top S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1+2\beta_1\boldsymbol{\gamma}_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\beta}_2+2\beta_1\beta_4\boldsymbol{\gamma}_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3)\\
&-\frac{1}{2(1-\rho_2^2)}(\beta_4^2\boldsymbol{\gamma}_1^\top S_4^{-1}RS_4^{-1}\boldsymbol{\gamma}_1+2\rho_2\beta_4\boldsymbol{\gamma}_1^\top S_2^{-1}S_4^{-1}\hat{\boldsymbol{\gamma}}_3-2\beta_4\boldsymbol{\gamma}_1^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3)\\
&-\frac{1}{2(1-\rho_2^2)}(2\beta_1\beta_4\boldsymbol{\gamma}_1^\top S_4^{-1}RS_4^{-1}\boldsymbol{\gamma}_3-2\rho_2\beta_4\boldsymbol{\gamma}_1^\top S_2^{-1}RS_4^{-1}\boldsymbol{\gamma}_3)-\frac{\gamma_{1j}^2}{2\sigma_1^2}\\
=&-\frac{1}{2}[\frac{(S_1^{-1}RS_1^{-1})_{jj}}{(1-\rho_1^2)}-\frac{2\rho_1\beta_1(S_3^{-1}RS_1^{-1})_{jj}}{(1-\rho_1^2)}+\frac{\beta_1^2(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)}+\frac{\beta_4^2(S_4^{-1}RS_4^{-1})_{jj}}{(1-\rho_2^2)}+\frac{1}{\sigma_1^2}]\gamma_{1j}^2\\
&-\frac{\sum_{j'\neq j}(S_1^{-1}RS_1^{-1})_{jj'}\gamma_{1j'}\gamma_{1j}}{(1-\rho_1^2)}+\frac{2\rho_1\beta_1\sum_{j'\neq j}(S_3^{-1}RS_1^{-1})_{jj'}\gamma_{1j'}\gamma_{1j}}{(1-\rho_1^2)}-\frac{\beta_1^2\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\gamma_{1j'}\gamma_{1j}}{(1-\rho_1^2)}\\
&-\frac{\rho_1\beta_1\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}\gamma_{1j}}{(1-\rho_1^2)}+\frac{\beta_1\sum_{j'}\hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}\gamma_{1j}}{(1-\rho_1^2)}\\
&+\frac{\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_1^{-1})_{j'j}\gamma_{1j}}{(1-\rho_1^2)}-\frac{\rho_1\sum_{j'}\hat{\Gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}\gamma_{1j}}{(1-\rho_1^2)}\\
&+\frac{\rho_1\sum_{j'}\beta_{2j'}(S_1^{-1}RS_3^{-1})_{j'j}\gamma_{1j}}{(1-\rho_1^2)}+\frac{\rho_1\beta_4\sum_{j'}\gamma_{3j'}(S_1^{-1}RS_3^{-1})_{j'j}\gamma_{1j}}{(1-\rho_1^2)}\\
&-\frac{\beta_1\sum_{j'}\beta_{2j'}(S_3^{-1}RS_3^{-1})_{j'j}\gamma_{1j}}{(1-\rho_1^2)}-\frac{\beta_1\beta_4\sum_{j'}\gamma_{3j'}(S_3^{-1}RS_3^{-1})_{j'j}\gamma_{1j}}{(1-\rho_1^2)}\\
&-\frac{\beta_4^2\sum_{j'\neq j}(S_4^{-1}RS_4^{-1})_{jj'}\gamma_{1j'}\gamma_{1j}}{(1-\rho_2^2)}-\frac{\rho_2\beta_4\sum_{j'}\hat{\gamma}_{3j'}(S_2^{-1}S_4^{-1})_{j'j}\gamma_{1j}}{(1-\rho_2^2)}+\frac{\beta_4\sum_{j'}\hat{\Gamma}_{3j'}(S_4^{-1}S_4^{-1})_{j'j}\gamma_{1j}}{(1-\rho_2^2)}\\
&-\frac{\beta_1\beta_4\sum_{j'}\gamma_{3j'}(S_4^{-1}RS_4^{-1})_{j'j}\gamma_{1j}}{(1-\rho_2^2)}+\frac{\rho_2\beta_4\sum_{j'}\gamma_{3j'}(S_2^{-1}RS_4^{-1})_{j'j}\gamma_{1j}}{(1-\rho_2^2)}.
\end{aligned}
$$

Let Ω denote all the latent variables. The conditional posterior distribution of each γ_{1j} given the

other parameters in the model is, $\gamma_{1j} \mid \Omega \backslash \gamma_{1j} \sim \mathcal{N}(\mu_{1j}, \sigma_{1j}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{1j}^2} =& -\frac{1}{2}[\frac{(S_1^{-1}RS_1^{-1})_{jj}}{(1-\rho_1^2)} - \frac{2\rho_1\beta_1(S_3^{-1}RS_1^{-1})_{jj}}{(1-\rho_1^2)} + \frac{\beta_1^2(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)} + \frac{\beta_4^2(S_4^{-1}RS_4^{-1})_{jj}}{(1-\rho_2^2)} + \frac{1}{\sigma_1^2}], \\
\frac{\mu_{1j}}{\sigma_{1j}^2} =& -\frac{\sum_{j'\neq j}(S_1^{-1}RS_1^{-1})_{jj'}\gamma_{1j'}}{(1-\rho_1^2)} + \frac{2\rho_1\beta_1\sum_{j'\neq j}(S_3^{-1}RS_1^{-1})_{jj'}\gamma_{1j'}}{(1-\rho_1^2)} - \frac{\beta_1^2\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\gamma_{1j'}}{(1-\rho_1^2)} \\
& - \frac{\rho_1\beta_1\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} + \frac{\beta_1\sum_{j'}\hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
& + \frac{\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_1^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\rho_1\sum_{j'}\hat{\Gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
& + \frac{\rho_1\sum_{j'}\beta_{2j'}(S_1^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} + \frac{\rho_1\beta_4\sum_{j'}\gamma_{3j'}(S_1^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
& - \frac{\beta_1\sum_{j'}\beta_{2j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_1\beta_4\sum_{j'}\gamma_{3j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
& - \frac{\beta_4^2\sum_{j'\neq j}(S_4^{-1}RS_4^{-1})_{jj'}\gamma_{1j'}}{(1-\rho_2^2)} - \frac{\rho_2\beta_4\sum_{j'}\hat{\gamma}_{3j'}(S_2^{-1}S_4^{-1})_{j'j}}{(1-\rho_2^2)} + \frac{\beta_4\sum_{j'}\hat{\Gamma}_{3j'}(S_4^{-1}S_4^{-1})_{j'j}}{(1-\rho_2^2)} \\
& - \frac{\beta_1\beta_4\sum_{j'}\gamma_{3j'}(S_4^{-1}RS_4^{-1})_{j'j}}{(1-\rho_2^2)} + \frac{\rho_2\beta_4\sum_{j'}\gamma_{3j'}(S_2^{-1}RS_4^{-1})_{j'j}}{(1-\rho_2^2)}.
\end{aligned}
$$

It follows that the full conditional distribution of γ_{3j} is

$$
\begin{aligned}
&\log\mathcal{N}(\begin{pmatrix}\hat{\boldsymbol{\gamma}}_1\\ \hat{\Gamma}_1\end{pmatrix};\begin{pmatrix}S_1RS_1^{-1}\boldsymbol{\gamma}_1\\ S_3RS_3^{-1}(\Gamma_1+\beta_4\boldsymbol{\gamma}_3)\end{pmatrix},\begin{pmatrix}S_1RS_1 & \rho_1S_1RS_3\\ \rho_1S_1RS_3 & S_3RS_3\end{pmatrix})\\
&+\log\mathcal{N}(\begin{pmatrix}\hat{\boldsymbol{\gamma}}_3\\ \hat{\Gamma}_3\end{pmatrix};\begin{pmatrix}S_2RS_2^{-1}\boldsymbol{\gamma}_3\\ S_4RS_4^{-1}\Gamma_3\end{pmatrix},\begin{pmatrix}S_2RS_2 & \rho_2S_2RS_4\\ \rho_2S_2RS_4 & S_4RS_4\end{pmatrix})\\
&+\log\mathcal{N}(\boldsymbol{\gamma}_1;\mathbf{0},\sigma_1^2\mathrm{I})+\log\mathcal{N}(\boldsymbol{\gamma}_3;\mathbf{0},\sigma_3^2\mathrm{I})\\
\propto&-\frac{1}{2(1-\rho_1^2)}(\boldsymbol{\gamma}_1^\top S_1^{-1}RS_1^{-1}\boldsymbol{\gamma}_1-2\boldsymbol{\gamma}_1^\top S_1^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1+2\rho_1(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1+2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\boldsymbol{\gamma}_1)\\
&-\frac{1}{2(1-\rho_1^2)}[-2\rho(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1-2(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}S_3^{-1}\hat{\Gamma}_1]\\
&-\frac{1}{2(1-\rho_1^2)}[(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}RS_3^{-1}(\Gamma_1+\beta_4\boldsymbol{\gamma}_3)]\\
&-\frac{1}{2(1-\rho_2^2)}(-2\boldsymbol{\gamma}_3^\top S_2^{-1}S_2^{-1}\hat{\boldsymbol{\gamma}}_3+2\rho_2\Gamma_3^\top S_2^{-1}S_4^{-1}\hat{\boldsymbol{\gamma}}_3+\boldsymbol{\gamma}_3^\top S_2^{-1}RS_2^{-1}\boldsymbol{\gamma}_3)\\
&-\frac{1}{2(1-\rho_2^2)}(2\rho_2\hat{\Gamma}_3^\top S_2^{-1}S_4^{-1}\boldsymbol{\gamma}_3-2\rho_2\Gamma_3^\top S_2^{-1}RS_4^{-1}\boldsymbol{\gamma}_3-2\Gamma_3^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3+\Gamma_3^\top S_4^{-1}RS_4^{-1}\Gamma_3)-\frac{\gamma_{3j}^2}{2\sigma_3^2}\\
\propto&-\frac{1}{2(1-\rho_1^2)}(\beta_4^2\boldsymbol{\gamma}_3^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3)-\frac{1}{2(1-\rho_1^2)}(2\rho_1\beta_4\boldsymbol{\gamma}_3^\top S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1-2\beta_4\boldsymbol{\gamma}_3^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1)\\
&-\frac{1}{2(1-\rho_1^2)}(-2\rho_1\beta_4\boldsymbol{\gamma}_3^\top S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1+2\beta_1\beta_4\boldsymbol{\gamma}_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3+2\beta_4\boldsymbol{\beta}_2^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3)\\
&-\frac{1}{2(1-\rho_2^2)}(\boldsymbol{\gamma}_3^\top S_2^{-1}RS_2^{-1}\boldsymbol{\gamma}_3-2\rho_2\beta_1\boldsymbol{\gamma}_3^\top S_2^{-1}RS_4^{-1}\boldsymbol{\gamma}_3+\beta_1^2\boldsymbol{\gamma}_3^\top S_4^{-1}RS_4^{-1}\boldsymbol{\gamma}_3)-\frac{\gamma_{3j}^2}{2\sigma_3^2}\\
&-\frac{1}{2(1-\rho_2^2)}(2\rho_2\beta_1\boldsymbol{\gamma}_3^\top S_2^{-1}S_4^{-1}\hat{\boldsymbol{\gamma}}_3-2\beta_1\boldsymbol{\gamma}_3^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3-2\boldsymbol{\gamma}_3^\top S_2^{-1}S_2^{-1}\hat{\boldsymbol{\gamma}}_3+2\rho_2\hat{\Gamma}_3^\top S_2^{-1}S_4^{-1}\boldsymbol{\gamma}_3)\\
&-\frac{1}{2(1-\rho_2^2)}(-2\rho_2\beta_4\boldsymbol{\gamma}_1^\top S_2^{-1}RS_4^{-1}\boldsymbol{\gamma}_3+2\beta_1\beta_4\boldsymbol{\gamma}_1^\top S_4^{-1}RS_4^{-1}\boldsymbol{\gamma}_3)\\
=&-\frac{1}{2}[\beta_4^2\frac{(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)}+\frac{(S_2^{-1}RS_2^{-1})_{jj}}{(1-\rho_2^2)}-\frac{2\rho_2\beta_1(S_2^{-1}RS_4^{-1})_{jj}}{(1-\rho_2^2)}+\frac{\beta_1^2(S_4^{-1}RS_4^{-1})_{jj}}{(1-\rho_2^2)}+\frac{1}{\sigma_3^2}]\gamma_{3j}^2\\
&-\frac{\beta_4^2\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\gamma_{3j'}\gamma_{3j}}{(1-\rho_1^2)}-\frac{\rho_1\beta_4\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)}+\frac{\beta_4\sum_{j'}\hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)}\\
&+\frac{\rho_1\beta_4\sum_{j'}\gamma_{1j'}(S_1^{-1}RS_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)}-\frac{\beta_1\beta_4\sum_{j'}\gamma_{1j'}(S_3^{-1}RS_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)}-\frac{\beta_4\sum_{j'}\beta_{2j'}(S_3^{-1}RS_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)}\\
&-\frac{\sum_{j'\neq j}(S_2^{-1}RS_2^{-1})_{jj'}\gamma_{3j'}\gamma_{3j}}{(1-\rho_2^2)}+\frac{2\rho_2\beta_1\sum_{j'\neq j}(S_2^{-1}RS_4^{-1})_{jj'}\gamma_{3j'}\gamma_{3j}}{(1-\rho_2^2)}-\frac{\beta_1^2\sum_{j'\neq j}(S_4^{-1}RS_4^{-1})_{jj'}\gamma_{3j'}\gamma_{3j}}{(1-\rho_2^2)}\\
&-\frac{\rho_2\beta_1\sum_{j'}\hat{\gamma}_{3j'}(S_2^{-1}S_4^{-1})_{j'j}\gamma_{3j}}{(1-\rho_2^2)}+\frac{\beta_1\sum_{j'}\hat{\Gamma}_{3j'}(S_4^{-1}S_4^{-1})_{j'j}\gamma_{3j}}{(1-\rho_2^2)}\\
&+\frac{\sum_{j'}\hat{\gamma}_{3j'}(S_2^{-1}S_2^{-1})_{j'j}\gamma_{3j}}{(1-\rho_2^2)}-\frac{\rho_2\sum_{j'}\hat{\Gamma}_{3j'}(S_2^{-1}S_4^{-1})_{j'j}\gamma_{3j}}{(1-\rho_2^2)}\\
&+\frac{\rho_2\beta_4\sum_{j'}\gamma_{1j'}(S_2^{-1}RS_4^{-1})_{j'j}\gamma_{3j}}{(1-\rho_2^2)}-\frac{\beta_1\beta_4\sum_{j'}\gamma_{1j'}(S_4^{-1}RS_4^{-1})_{j'j}\gamma_{3j}}{(1-\rho_2^2)}.
\end{aligned}
$$

The conditional posterior distribution is, $\gamma_{3j} \mid \Omega \backslash \gamma_{3j} \sim \mathcal{N}(\mu_{3j}, \sigma_{3j}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{3j}^2} =& -\frac{1}{2}[\beta_4^2 \frac{(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)} + \frac{(S_2^{-1}RS_2^{-1})_{jj}}{(1-\rho_2^2)} - \frac{2\rho_2\beta_1(S_2^{-1}RS_4^{-1})_{jj}}{(1-\rho_2^2)} + \frac{\beta_1^2(S_4^{-1}RS_4^{-1})_{jj}}{(1-\rho_2^2)} + \frac{1}{\sigma_3^2}], \\
\frac{\mu_{3j}}{\sigma_{3j}^2} =& -\frac{\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\gamma_{3j'}}{(1-\rho_1^2)} - \frac{\rho_1\beta_4\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} + \frac{\beta_4\sum_{j'}\hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
&+ \frac{\rho_1\beta_4\sum_{j'}\gamma_{1j'}(S_1^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_1\beta_4\sum_{j'}\gamma_{1j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_4\sum_{j'}\beta_{2j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
&- \frac{\sum_{j'\neq j}(S_2^{-1}RS_2^{-1})_{jj'}\gamma_{3j'}}{(1-\rho_2^2)} + \frac{2\rho_2\beta_1\sum_{j'\neq j}(S_2^{-1}RS_4^{-1})_{jj'}\gamma_{3j'}}{(1-\rho_2^2)} - \frac{\beta_1^2\sum_{j'\neq j}(S_4^{-1}RS_4^{-1})_{jj'}\gamma_{3j'}}{(1-\rho_2^2)} \\
&- \frac{\rho_2\beta_1\sum_{j'}\hat{\gamma}_{3j'}(S_2^{-1}S_4^{-1})_{j'j}}{(1-\rho_2^2)} + \frac{\beta_1\sum_{j'}\hat{\Gamma}_{3j'}(S_4^{-1}S_4^{-1})_{j'j}}{(1-\rho_2^2)} \\
&+ \frac{\sum_{j'}\hat{\gamma}_{3j'}(S_2^{-1}S_2^{-1})_{j'j}}{(1-\rho_2^2)} - \frac{\rho_2\sum_{j'}\hat{\Gamma}_{3j'}(S_2^{-1}S_4^{-1})_{j'j}}{(1-\rho_2^2)} \\
&+ \frac{\rho_2\beta_4\sum_{j'}\gamma_{1j'}(S_2^{-1}RS_4^{-1})_{j'j}}{(1-\rho_2^2)} - \frac{\beta_1\beta_4\sum_{j'}\gamma_{1j'}(S_4^{-1}RS_4^{-1})_{j'j}}{(1-\rho_2^2)}.
\end{aligned}
$$

The full conditional distribution of β_1 is

$$
\begin{aligned}
&\log\mathcal{N}(\begin{pmatrix}\hat{\gamma}_1\\ \hat{\Gamma}_1\end{pmatrix}; \begin{pmatrix}S_1RS_1^{-1}\gamma_1\\ S_3RS_3^{-1}(\Gamma_1+\beta_4\gamma_3)\end{pmatrix}, \begin{pmatrix}S_1RS_1 & \rho_1S_1RS_3\\ \rho_1S_1RS_3 & S_3RS_3\end{pmatrix}) \\
&+\log\mathcal{N}(\begin{pmatrix}\hat{\gamma}_3\\ \hat{\Gamma}_3\end{pmatrix}; \begin{pmatrix}S_2RS_2^{-1}\gamma_3\\ S_4RS_4^{-1}\Gamma_3\end{pmatrix}, \begin{pmatrix}S_2RS_2 & \rho_2S_2RS_4\\ \rho_2S_2RS_4 & S_4RS_4\end{pmatrix}) + \log\Pr(\beta_1) \\
\propto& -\frac{1}{2(1-\rho_1^2)}[\gamma_1^\top S_1^{-1}RS_1^{-1}\gamma_1 - 2\gamma_1^\top S_1^{-1}S_1^{-1}\hat{\gamma}_1 + 2\rho_1(\Gamma_1^\top+\beta_4\gamma_3^\top)S_3^{-1}S_1^{-1}\hat{\gamma}_1 + 2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\gamma_1] \\
&-\frac{1}{2(1-\rho_1^2)}[-2\rho(\Gamma_1^\top+\beta_4\gamma_3^\top)S_3^{-1}RS_1^{-1}\gamma_1 - 2(\Gamma_1^\top+\beta_4\gamma_3^\top)S_3^{-1}S_3^{-1}\hat{\Gamma}_1] \\
&-\frac{1}{2(1-\rho_1^2)}[(\Gamma_1^\top+\beta_4\gamma_3^\top)S_3^{-1}RS_3^{-1}(\Gamma_1+\beta_4\gamma_3)] \\
&-\frac{1}{2(1-\rho_2^2)}(-2\gamma_3^\top S_2^{-1}S_2^{-1}\hat{\gamma}_3 + 2\rho_2\Gamma_3^\top S_2^{-1}S_4^{-1}\hat{\gamma}_3 + \gamma_3^\top S_2^{-1}RS_2^{-1}\gamma_3) \\
&-\frac{1}{2(1-\rho_2^2)}(2\rho_2\hat{\Gamma}_3^\top S_2^{-1}S_4^{-1}\gamma_3 - 2\rho_2\Gamma_3^\top S_2^{-1}RS_4^{-1}\gamma_3 - 2\Gamma_3^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3 + \Gamma_3^\top S_4^{-1}RS_4^{-1}\Gamma_3) \\
\propto& -\frac{1}{2}[\frac{\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_1}{(1-\rho_1^2)} + \frac{\gamma_3^\top S_4^{-1}RS_4^{-1}\gamma_3}{(1-\rho_2^2)}]\beta_1^2 \\
&-\frac{\rho_1\gamma_1^\top S_3^{-1}S_1^{-1}\hat{\gamma}_1 - \gamma_1^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1}{(1-\rho_1^2)}\beta_1 \\
&-\frac{-\rho_1\gamma_1^\top S_3^{-1}RS_1^{-1}\gamma_1 + \gamma_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\beta}_2 + \beta_4\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_3}{(1-\rho_1^2)}\beta_1 \\
&-\frac{\rho_2\gamma_3^\top S_2^{-1}S_4^{-1}\hat{\gamma}_3 - \gamma_3^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3}{(1-\rho_2^2)}\beta_1 \\
&-\frac{-\rho_2\gamma_3^\top S_2^{-1}RS_4^{-1}\gamma_3 + \beta_4\gamma_1^\top S_4^{-1}RS_4^{-1}\gamma_3}{(1-\rho_2^2)}\beta_1.
\end{aligned}
$$

The conditional posterior distribution is, $\beta_1 \mid \Omega \backslash \beta_1 \sim \mathcal{N}(\mu_{b1}, \sigma_{b1}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{b1}^2} =& -\frac{1}{2}[\frac{\boldsymbol{\gamma}_1^\top S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}_1}{(1-\rho_1^2)} + \frac{\boldsymbol{\gamma}_3^\top S_4^{-1} R S_4^{-1} \boldsymbol{\gamma}_3}{(1-\rho_2^2)}], \\
\frac{\mu_{b1}}{\sigma_{b1}^2} =& -\frac{\rho_1 \boldsymbol{\gamma}_1^\top S_3^{-1} S_1^{-1} \hat{\boldsymbol{\gamma}}_1 - \boldsymbol{\gamma}_1^\top S_3^{-1} S_3^{-1} \hat{\Gamma}_1}{(1-\rho_1^2)} \\
& - \frac{-\rho_1 \boldsymbol{\gamma}_1^\top S_3^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_1^\top S_3^{-1} R S_3^{-1} \boldsymbol{\beta}_2 + \beta_4 \boldsymbol{\gamma}_1^\top S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}_3}{(1-\rho_1^2)} \\
& - \frac{\rho_2 \boldsymbol{\gamma}_3^\top S_2^{-1} S_4^{-1} \hat{\boldsymbol{\gamma}}_3 - \boldsymbol{\gamma}_3^\top S_4^{-1} S_4^{-1} \hat{\Gamma}_3}{(1-\rho_2^2)} \\
& - \frac{-\rho_2 \boldsymbol{\gamma}_3^\top S_2^{-1} R S_4^{-1} \boldsymbol{\gamma}_3 + \beta_4 \boldsymbol{\gamma}_1^\top S_4^{-1} R S_4^{-1} \boldsymbol{\gamma}_3}{(1-\rho_2^2)}.
\end{aligned}
$$

The full conditional distribution of β_{2j} is

$$
\begin{aligned}
& \log \mathcal{N}(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_1 \\ \hat{\Gamma}_1 \end{pmatrix}; \begin{pmatrix} S_1 R S_1^{-1} \boldsymbol{\gamma}_1 \\ S_3 R S_3^{-1} (\Gamma_1 + \beta_4 \boldsymbol{\gamma}_3) \end{pmatrix}, \begin{pmatrix} S_1 R S_1 & \rho_1 S_1 R S_3 \\ \rho_1 S_1 R S_3 & S_3 R S_3 \end{pmatrix}) + \log \mathcal{N}(\boldsymbol{\beta}_2; \mathbf{0}, \sigma_2^2 \mathrm{I}) \\
\propto & -\frac{1}{2(1-\rho_1^2)}[\boldsymbol{\gamma}_1^\top S_1^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 - 2\boldsymbol{\gamma}_1^\top S_1^{-1} S_1^{-1} \hat{\boldsymbol{\gamma}}_1 + 2\rho_1 (\Gamma_1^\top + \beta_4 \boldsymbol{\gamma}_3^\top) S_3^{-1} S_1^{-1} \hat{\boldsymbol{\gamma}}_1 + 2\rho_1 \hat{\Gamma}_1^\top S_3^{-1} S_1^{-1} \boldsymbol{\gamma}_1] \\
& - \frac{1}{2(1-\rho_1^2)}[-2\rho (\Gamma_1^\top + \beta_4 \boldsymbol{\gamma}_3^\top) S_3^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 - 2(\Gamma_1^\top + \beta_4 \boldsymbol{\gamma}_3^\top) S_3^{-1} S_3^{-1} \hat{\Gamma}_1] \\
& - \frac{1}{2(1-\rho_1^2)}[(\Gamma_1^\top + \beta_4 \boldsymbol{\gamma}_3^\top) S_3^{-1} R S_3^{-1} (\Gamma_1 + \beta_4 \boldsymbol{\gamma}_3)] \\
& - \frac{\beta_{2j}^2}{2\sigma_2^2} \\
\propto & -\frac{1}{2(1-\rho_1^2)}(\boldsymbol{\beta}_2^\top S_3^{-1} R S_3^{-1} \boldsymbol{\beta}_2) \\
& - \frac{1}{2(1-\rho_1^2)}(2\rho_1 \boldsymbol{\beta}_2^\top S_3^{-1} S_1^{-1} \hat{\boldsymbol{\gamma}}_1 - 2\boldsymbol{\beta}_2^\top S_3^{-1} S_3^{-1} \hat{\Gamma}_1) \\
& - \frac{1}{2(1-\rho_1^2)}(-2\rho_1 \boldsymbol{\gamma}_1^\top S_3^{-1} R S_1^{-1} \boldsymbol{\beta}_2 + 2\beta_1 \boldsymbol{\gamma}_1^\top S_3^{-1} R S_3^{-1} \boldsymbol{\beta}_2 + 2\beta_4 \boldsymbol{\beta}_2^\top S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}_3) \\
& - \frac{\beta_{2j}^2}{2\sigma_2^2} \\
= & -\frac{1}{2}[\frac{(S_3^{-1} R S_3^{-1})_{jj}}{(1-\rho_1^2)} + \frac{1}{\sigma_2^2}]\beta_{2j}^2 \\
& - \frac{\sum_{j' \neq j} (S_3^{-1} R S_3^{-1})_{jj'} \beta_{2j'} \beta_{2j}}{(1-\rho_1^2)} \\
& - \frac{\rho_1 \sum_{j'} \hat{\gamma}_{1j'} (S_1^{-1} S_3^{-1})_{j'j} \beta_{2j}}{(1-\rho_1^2)} + \frac{\sum_{j'} \hat{\Gamma}_{1j'} (S_3^{-1} S_3^{-1})_{j'j} \beta_{2j}}{(1-\rho_1^2)} \\
& + \frac{\rho_1 \sum_{j'} \gamma_{1j'} (S_1^{-1} R S_3^{-1})_{j'j} \beta_{2j}}{(1-\rho_1^2)} - \frac{\beta_1 \sum_{j'} \gamma_{1j'} (S_3^{-1} R S_3^{-1})_{j'j} \beta_{2j}}{(1-\rho_1^2)} - \frac{\beta_4 \sum_{j'} \gamma_{3j'} (S_3^{-1} R S_3^{-1})_{j'j} \beta_{2j}}{(1-\rho_1^2)}.
\end{aligned}
$$

The conditional posterior distribution is, $\beta_{2j} \mid \Omega \backslash \beta_{2j} \sim \mathcal{N}(\mu_{b2}, \sigma_{b2}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{b2}^2} =& -\frac{1}{2}\left(\frac{(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)} + \frac{1}{\sigma_2^2}\right), \\
\frac{\mu_{b2}}{\sigma_{b2}^2} =& -\frac{\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\beta_{2j'}}{(1-\rho_1^2)} \\
& - \frac{\rho_1 \sum_{j'} \hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} + \frac{\sum_{j'} \hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
& + \frac{\rho_1 \sum_{j'} \gamma_{1j'}(S_1^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_1 \sum_{j'} \gamma_{1j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_4 \sum_{j'} \gamma_{3j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)}.
\end{aligned}
$$

The full conditional distribution of β_4 is

$$
\begin{aligned}
& \log \mathcal{N}\left(\begin{pmatrix}\hat{\gamma}_1 \\ \hat{\Gamma}_1\end{pmatrix}; \begin{pmatrix}S_1RS_1^{-1}\gamma_1 \\ S_3RS_3^{-1}(\Gamma_1+\beta_4\gamma_3)\end{pmatrix}, \begin{pmatrix}S_1RS_1 & \rho_1 S_1RS_3 \\ \rho_1 S_1RS_3 & S_3RS_3\end{pmatrix}\right) \\
& + \log \mathcal{N}\left(\begin{pmatrix}\hat{\gamma}_3 \\ \hat{\Gamma}_3\end{pmatrix}; \begin{pmatrix}S_2RS_2^{-1}\gamma_3 \\ S_4RS_4^{-1}\Gamma_3\end{pmatrix}, \begin{pmatrix}S_2RS_2 & \rho_2 S_2RS_4 \\ \rho_2 S_2RS_4 & S_4RS_4\end{pmatrix}\right) + \log \Pr(\beta_1) \\
\propto & -\frac{1}{2(1-\rho_1^2)}[\gamma_1^\top S_1^{-1}RS_1^{-1}\gamma_1 - 2\gamma_1^\top S_1^{-1}S_1^{-1}\hat{\gamma}_1 + 2\rho_1(\Gamma_1^\top + \beta_4\gamma_3^\top)S_3^{-1}S_1^{-1}\hat{\gamma}_1 + 2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\gamma_1] \\
& -\frac{1}{2(1-\rho_1^2)}[-2\rho(\Gamma_1^\top + \beta_4\gamma_3^\top)S_3^{-1}RS_1^{-1}\gamma_1 - 2(\Gamma_1^\top + \beta_4\gamma_3^\top)S_3^{-1}S_3^{-1}\hat{\Gamma}_1] \\
& -\frac{1}{2(1-\rho_1^2)}[(\Gamma_1^\top + \beta_4\gamma_3^\top)S_3^{-1}RS_3^{-1}(\Gamma_1+\beta_4\gamma_3)] \\
& -\frac{1}{2(1-\rho_2^2)}(-2\gamma_3^\top S_2^{-1}S_2^{-1}\hat{\gamma}_3 + 2\rho_2\Gamma_3^\top S_2^{-1}S_4^{-1}\hat{\gamma}_3 + \gamma_3^\top S_2^{-1}RS_2^{-1}\gamma_3) \\
& -\frac{1}{2(1-\rho_2^2)}(2\rho_2\hat{\Gamma}_3^\top S_2^{-1}S_4^{-1}\gamma_3 - 2\rho_2\Gamma_3^\top S_2^{-1}RS_4^{-1}\gamma_3 - 2\Gamma_3^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3 + \Gamma_3^\top S_4^{-1}RS_4^{-1}\Gamma_3) \\
\propto & -\frac{1}{2}\left[\frac{\gamma_3^\top S_3^{-1}RS_3^{-1}\gamma_3}{(1-\rho_1^2)} + \frac{\gamma_1^\top S_4^{-1}RS_4^{-1}\gamma_1}{(1-\rho_2^2)}\right]\beta_4^2 \\
& - \frac{\rho_1\gamma_3^\top S_3^{-1}S_1^{-1}\hat{\gamma}_1 - \gamma_3^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1}{(1-\rho_1^2)}\beta_4 \\
& - \frac{-\rho_1\gamma_3^\top S_3^{-1}RS_1^{-1}\gamma_1 + \beta_1\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_3 + \boldsymbol{\beta}_2^\top S_3^{-1}RS_3^{-1}\gamma_3}{(1-\rho_1^2)}\beta_4 \\
& - \frac{\rho_2\gamma_1^\top S_2^{-1}S_4^{-1}\hat{\gamma}_3 - \gamma_1^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3}{(1-\rho_2^2)}\beta_4 \\
& - \frac{-\rho_2\gamma_1^\top S_2^{-1}RS_4^{-1}\gamma_3 + \beta_1\gamma_1^\top S_4^{-1}RS_4^{-1}\gamma_3}{(1-\rho_2^2)}\beta_4.
\end{aligned}
$$

The conditional posterior distribution is, $\beta_4 \mid \Omega \backslash \beta_4 \sim \mathcal{N}(\mu_{b4}, \sigma_{b4}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{b4}^2} =& -\frac{1}{2}\left[\frac{\gamma_3^\top S_3^{-1}RS_3^{-1}\gamma_3}{(1-\rho_1^2)} + \frac{\gamma_1^\top S_4^{-1}RS_4^{-1}\gamma_1}{(1-\rho_2^2)}\right], \\
\frac{\mu_{b4}}{\sigma_{b4}^2} =& -\frac{\rho_1\gamma_3^\top S_3^{-1}S_1^{-1}\hat{\gamma}_1 - \gamma_3^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1}{(1-\rho_1^2)} \\
& - \frac{-\rho_1\gamma_3^\top S_3^{-1}RS_1^{-1}\gamma_1 + \beta_1\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_3 + \boldsymbol{\beta}_2^\top S_3^{-1}RS_3^{-1}\gamma_3}{(1-\rho_1^2)} \\
& - \frac{\rho_2\gamma_1^\top S_2^{-1}S_4^{-1}\hat{\gamma}_3 - \gamma_1^\top S_4^{-1}S_4^{-1}\hat{\Gamma}_3}{(1-\rho_2^2)} \\
& - \frac{-\rho_2\gamma_1^\top S_2^{-1}RS_4^{-1}\gamma_3 + \beta_1\gamma_1^\top S_4^{-1}RS_4^{-1}\gamma_3}{(1-\rho_2^2)}.
\end{aligned}
$$

The full conditional distribution of σ_1^2 is

$$\begin{aligned}
&\log \Pr(\boldsymbol{\gamma}|\sigma_1^2, \sigma_3^2) + \log \Pr(\sigma_1^2) \\
\propto &- \frac{\boldsymbol{\gamma}_1^\top \boldsymbol{\gamma}_1}{2\sigma_1^2} - (a_1 + 1) \log \sigma_1^2 - \frac{b_1}{\sigma_1^2} - \frac{p}{2} \log \sigma_1^2 \\
\propto &- \frac{\boldsymbol{\gamma}_1^\top \boldsymbol{\gamma}_1 + 2b_1}{2\sigma_1^2} - (a_1 + \frac{p}{2} + 1) \log \sigma_1^2.
\end{aligned}$$

So the posterior distribution of σ_1^2 is

$$\sigma_1^2 \mid \Omega \backslash \sigma_1^2 \sim \text{Inv-Gamma}(a_1 + \frac{p}{2}, b_1 + \frac{\boldsymbol{\gamma}_1^\top \boldsymbol{\gamma}_1}{2})$$

The full conditional distribution of σ_3^2 is

$$\begin{aligned}
&\log Pr(\boldsymbol{\gamma}|\sigma_1^2, \sigma_3^2) + \log \Pr(\sigma_3^2) \\
\propto &- \frac{\boldsymbol{\gamma}_3^\top \boldsymbol{\gamma}_3}{2\sigma_3^2} - (a_3 + 1) \log \sigma_3^2 - \frac{b_3}{\sigma_3^2} - \frac{p}{2} \log \sigma_3^2 \\
\propto &- \frac{\boldsymbol{\gamma}_3^\top \boldsymbol{\gamma}_3 + 2b_3}{2\sigma_3^2} - (a_3 + \frac{p}{2} + 1) \log \sigma_3^2.
\end{aligned}$$

So the posterior distribution of σ_3^2 is

$$\sigma_3^2 \mid \Omega \backslash \sigma_3^2 \sim \text{Inv-Gamma}(a_3 + \frac{p}{2}, b_3 + \frac{\boldsymbol{\gamma}_3^\top \boldsymbol{\gamma}_3}{2})$$

Similarly, the posterior distribution of σ_2^2 is

$$\sigma_2^2 \mid \Omega \backslash \sigma_2^2 \sim \text{Inv-Gamma}(a_2 + \frac{p}{2}, b_2 + \frac{\boldsymbol{\beta}_2^\top \boldsymbol{\beta}_2}{2}).$$

S3.2 Gibbs sampler for MERLIN without sample overlap

If there is no sample overlap, the correlation between the exposure and outcome variables due to that overlap is zero; thus, $\rho_1 = \rho_2 = 0$.

S3.3 Gibbs sampler for MERLIN with an unbalanced binary E

The full log-likelihood function is

$$\begin{aligned}
L(\theta|\mathcal{D}) = &\log \mathcal{N}(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_1 \\ \hat{\Gamma}_1 \end{pmatrix}; \begin{pmatrix} S_1 R S_1^{-1} \boldsymbol{\gamma}_1 \\ S_3 R S_3^{-1} (\Gamma_1 + \beta_4 \boldsymbol{\gamma}_3) \end{pmatrix}, \begin{pmatrix} S_1 R S_1 & \rho_1 S_1 R S_3 \\ \rho_1 S_1 R S_3 & S_3 R S_3 \end{pmatrix}) \\
&+ \log \mathcal{N}(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_3 \\ \hat{\Gamma}_3 \end{pmatrix}; \begin{pmatrix} S_2 R S_2^{-1} \boldsymbol{\gamma}_3 \\ S_4 R S_4^{-1} \Gamma_3 + \mu_3 \beta_4 S_4^2 S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}_3 \end{pmatrix}, \begin{pmatrix} S_2 R S_2 & \rho_2 S_2 R S_4 \\ \rho_2 S_2 R S_4 & S_4 R S_4 \end{pmatrix}) \\
&+ \log \mathcal{N}(\boldsymbol{\gamma}_1; \mathbf{0}, \sigma_1^2 \mathrm{I}) + \log \mathcal{N}(\boldsymbol{\beta}_2; 0, \sigma_2^2 \mathrm{I}) + \log \mathcal{N}(\boldsymbol{\gamma}_3; \mathbf{0}, \sigma_3^2 \mathrm{I})
\end{aligned} \tag{9}$$

The Gibbs sampler can be derived using a similar approach.

S3.4 Gibbs sampler for MERLIN with a continuous E

The full log-likelihood function is

$$\begin{aligned}
L(\theta|\mathcal{D}) = &\log \mathcal{N}(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_1 \\ \hat{\Gamma}_1 \end{pmatrix}; \begin{pmatrix} S_1 R S_1^{-1} \boldsymbol{\gamma}_1 \\ S_3 R S_3^{-1} \Gamma_1 + S_3^2 S_4^{-1} R S_4^{-1} \beta_4 \boldsymbol{\gamma}_3 \end{pmatrix}, \begin{pmatrix} S_1 R S_1 & \rho_1 S_1 R S_3 \\ \rho_1 S_1 R S_3 & S_3 R S_3 \end{pmatrix}) \\
&+ \log \mathcal{N}(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_3 \\ \hat{\Gamma}_3 \end{pmatrix}; \begin{pmatrix} S_2 R S_2^{-1} \boldsymbol{\gamma}_3 \\ S_4 R S_4^{-1} \Gamma_3 \end{pmatrix}, \begin{pmatrix} S_2 R S_2 & \rho_2 S_2 R S_4 \\ \rho_2 S_2 R S_4 & S_4 R S_4 \end{pmatrix}) \\
&+ \log \mathcal{N}(\boldsymbol{\gamma}_1; \mathbf{0}, \sigma_1^2 \mathrm{I}) + \log \mathcal{N}(\boldsymbol{\beta}_2; 0, \sigma_2^2 \mathrm{I}) + \log \mathcal{N}(\boldsymbol{\gamma}_3; \mathbf{0}, \sigma_3^2 \mathrm{I})
\end{aligned} \tag{10}$$

The Gibbs sampler can be derived using a similar approach, too.

S4 Parameter Estimation and Statistical Inference for MERLIN(p)

When phenotypic outcome data are difficult to obtain for a particular GWIS, we provide MERLIN(part) algorithms that still require exposure-GWIS data but can be used without outcome-GWIS data.While yielding unbiased estimates and controlled Type I error for β_1 and β_4, MERLIN(part) has less power than full MERLIN (Supplementary Fig. S14, Supplementary Notes as below for model details). The same as MERLIN, we set $\boldsymbol{\gamma}^{(G)} = \boldsymbol{\gamma}_1, \boldsymbol{\gamma}^{(GI)} = \boldsymbol{\gamma}_3, \Gamma^{(G)} = \Gamma_1, \Gamma^{(GI)} = \Gamma_3, \beta^{(A)} = \beta_1, \beta^{(I)} = \beta_4, \boldsymbol{\beta}^{(G)} = \boldsymbol{\beta}_2$.

The distribution of the available summary statistics is

$$\begin{pmatrix} \hat{\boldsymbol{\gamma}}_1 \\ \hat{\boldsymbol{\gamma}}_3 \\ \hat{\Gamma}_1 \end{pmatrix} \mid \boldsymbol{\gamma} \overset{A}{\sim} \mathcal{N}\left(\begin{pmatrix} S_1 R S_1^{-1} \boldsymbol{\gamma}_1 \\ S_2 R S_2^{-1} \boldsymbol{\gamma}_3 \\ S_3 R S_3^{-1} (\Gamma_1 + \beta_4 \boldsymbol{\gamma}_3) \end{pmatrix}, \begin{bmatrix} S_1 R S_1 & \mathbf{0} & \rho_1 S_1 R S_3 \\ \mathbf{0} & S_2 R S_2 & \mathbf{0} \\ \rho_1 S_1 R S_3 & \mathbf{0} & S_3 R S_3 \end{bmatrix} \right).$$

It follows that the full conditional distribution of γ_{1j} is

$$\begin{aligned}
&\log \mathcal{N}\left(\begin{pmatrix} \hat{\boldsymbol{\gamma}}_1 \\ \hat{\Gamma}_1 \end{pmatrix}; \begin{pmatrix} S_1 R S_1^{-1} \boldsymbol{\gamma}_1 \\ S_3 R S_3^{-1} (\Gamma_1 + \beta_4 \boldsymbol{\gamma}_3) \end{pmatrix}, \begin{pmatrix} S_1 R S_1 & \rho_1 S_1 R S_3 \\ \rho_1 S_1 R S_3 & S_3 R S_3 \end{pmatrix}\right) + \log \mathcal{N}(\boldsymbol{\gamma}_1; \mathbf{0}, \sigma_1^2 \mathrm{I}) + \log \mathcal{N}(\boldsymbol{\gamma}_3; \mathbf{0}, \sigma_3^2 \mathrm{I}) \\
\propto & -\frac{1}{2(1-\rho_1^2)} (\boldsymbol{\gamma}_1^\top S_1^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 - 2\boldsymbol{\gamma}_1^\top S_1^{-1} S_1^{-1} \hat{\boldsymbol{\gamma}}_1 + 2\rho_1 (\Gamma_1^\top + \beta_4 \boldsymbol{\gamma}_3^\top) S_3^{-1} S_1^{-1} \hat{\boldsymbol{\gamma}}_1 + 2\rho_1 \hat{\Gamma}_1^\top S_3^{-1} S_1^{-1} \boldsymbol{\gamma}_1) \\
& - \frac{1}{2(1-\rho_1^2)} (-2\rho (\Gamma_1^\top + \beta_4 \boldsymbol{\gamma}_3^\top) S_3^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 - 2(\Gamma_1^\top + \beta_4 \boldsymbol{\gamma}_3^\top) S_3^{-1} S_3^{-1} \hat{\Gamma}_1 + (\Gamma_1^\top + \beta_4 \boldsymbol{\gamma}_3^\top) S_3^{-1} R S_3^{-1} (\Gamma_1 + \beta_4 \boldsymbol{\gamma}_3)) \\
& - \frac{\gamma_{1j}^2}{2\sigma_1^2} \\
\propto & -\frac{1}{2(1-\rho_1^2)} (\boldsymbol{\gamma}_1^\top S_1^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 - 2\rho_1 \beta_1 \boldsymbol{\gamma}_1^\top S_3^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 + \beta_1^2 \boldsymbol{\gamma}_1^\top S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}_1) \\
& - \frac{1}{2(1-\rho_1^2)} (2\rho_1 \beta_1 \boldsymbol{\gamma}_1^\top S_3^{-1} S_1^{-1} \hat{\boldsymbol{\gamma}}_1 - 2\beta_1 \boldsymbol{\gamma}_1^\top S_3^{-1} S_3^{-1} \hat{\Gamma}_1 - 2\boldsymbol{\gamma}_1^\top S_1^{-1} S_1^{-1} \hat{\boldsymbol{\gamma}}_1 + 2\rho_1 \hat{\Gamma}_1^\top S_3^{-1} S_1^{-1} \boldsymbol{\gamma}_1) \\
& - \frac{1}{2(1-\rho_1^2)} (-2\rho_1 \boldsymbol{\beta}_2^\top S_3^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 - 2\rho_1 \beta_4 \boldsymbol{\gamma}_3^\top S_3^{-1} R S_1^{-1} \boldsymbol{\gamma}_1 + 2\beta_1 \boldsymbol{\gamma}_1^\top S_3^{-1} R S_3^{-1} \boldsymbol{\beta}_2 + 2\beta_1 \beta_4 \boldsymbol{\gamma}_1^\top S_3^{-1} R S_3^{-1} \boldsymbol{\gamma}_3) \\
& - \frac{\gamma_{1j}^2}{2\sigma_1^2} \\
= & -\frac{1}{2} \Big(\frac{(S_1^{-1} R S_1^{-1})_{jj}}{(1-\rho_1^2)} - \frac{2\rho_1 \beta_1 (S_3^{-1} R S_1^{-1})_{jj}}{(1-\rho_1^2)} + \frac{\beta_1^2 (S_3^{-1} R S_3^{-1})_{jj}}{(1-\rho_1^2)} + \frac{1}{\sigma_1^2} \Big) \gamma_{1j}^2 \\
& - \frac{\sum_{j' \neq j} (S_1^{-1} R S_1^{-1})_{jj'} \gamma_{1j'} \gamma_{1j}}{(1-\rho_1^2)} + \frac{2\rho_1 \beta_1 \sum_{j' \neq j} (S_3^{-1} R S_1^{-1})_{jj'} \gamma_{1j'} \gamma_{1j}}{(1-\rho_1^2)} - \frac{\beta_1^2 \sum_{j' \neq j} (S_3^{-1} R S_3^{-1})_{jj'} \gamma_{1j'} \gamma_{1j}}{(1-\rho_1^2)} \\
& - \frac{\rho_1 \beta_1 \sum_{j'} \hat{\gamma}_{1j'} (S_1^{-1} S_3^{-1})_{j'j} \gamma_{1j}}{(1-\rho_1^2)} + \frac{\beta_1 \sum_{j'} \hat{\Gamma}_{1j'} (S_3^{-1} S_3^{-1})_{j'j} \gamma_{1j}}{(1-\rho_1^2)} \\
& + \frac{\sum_{j'} \hat{\gamma}_{1j'} (S_1^{-1} S_1^{-1})_{j'j} \gamma_{1j}}{(1-\rho_1^2)} - \frac{\rho_1 \sum_{j'} \hat{\Gamma}_{1j'} (S_1^{-1} S_3^{-1})_{j'j} \gamma_{1j}}{(1-\rho_1^2)} \\
& + \frac{\rho_1 \sum_{j'} \beta_{2j'} (S_1^{-1} R S_3^{-1})_{j'j} \gamma_{1j}}{(1-\rho_1^2)} + \frac{\rho_1 \beta_4 \sum_{j'} \gamma_{3j'} (S_1^{-1} R S_3^{-1})_{j'j} \gamma_{1j}}{(1-\rho_1^2)} \\
& - \frac{\beta_1 \sum_{j'} \beta_{2j'} (S_3^{-1} R S_3^{-1})_{j'j} \gamma_{1j}}{(1-\rho_1^2)} - \frac{\beta_1 \beta_4 \sum_{j'} \gamma_{3j'} (S_3^{-1} R S_3^{-1})_{j'j} \gamma_{1j}}{(1-\rho_1^2)}.
\end{aligned}$$

Let Ω denote all the latent variables. Obviously, the conditional posterior distribution is $\gamma_{1j} \mid$

$\Omega\backslash\gamma_{1j} \sim \mathcal{N}(\mu_{1j}, \sigma_{1j}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{1j}^2} =& -\frac{1}{2}\Big(\frac{(S_1^{-1}RS_1^{-1})_{jj}}{(1-\rho_1^2)} - \frac{2\rho_1\beta_1(S_3^{-1}RS_1^{-1})_{jj}}{(1-\rho_1^2)} + \frac{\beta_1^2(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)} + \frac{1}{\sigma_1^2}\Big),\\
\frac{\mu_{1j}}{\sigma_{1j}^2} =& -\frac{\sum_{j'\neq j}(S_1^{-1}RS_1^{-1})_{jj'}\gamma_{1j'}}{(1-\rho_1^2)} + \frac{2\rho_1\beta_1\sum_{j'\neq j}(S_3^{-1}RS_1^{-1})_{jj'}\gamma_{1j'}}{(1-\rho_1^2)} - \frac{\beta_1^2\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\gamma_{1j'}}{(1-\rho_1^2)}\\
&- \frac{\rho_1\beta_1\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} + \frac{\beta_1\sum_{j'}\hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)}\\
&+ \frac{\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_1^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\rho_1\sum_{j'}\hat{\Gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)}\\
&+ \frac{\rho_1\sum_{j'}\beta_{2j'}(S_1^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} + \frac{\rho_1\beta_4\sum_{j'}\gamma_{3j'}(S_1^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)}\\
&- \frac{\beta_1\sum_{j'}\beta_{2j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_1\beta_4\sum_{j'}\gamma_{3j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)}.
\end{aligned}
$$

It follows that the full conditional distribution of γ_{3j} is

$$
\begin{aligned}
&\log\mathcal{N}\Big(\begin{pmatrix}\hat{\gamma}_1\\ \hat{\Gamma}_1\end{pmatrix}; \begin{pmatrix}S_1RS_1^{-1}\gamma_1\\ S_3RS_3^{-1}(\Gamma_1+\beta_4\gamma_3)\end{pmatrix}, \begin{pmatrix}S_1RS_1 & \rho_1S_1RS_3\\ \rho_1S_1RS_3 & S_3RS_3\end{pmatrix}\Big) + \log\mathcal{N}(\hat{\gamma}_3; S_2RS_2^{-1}\gamma_3, S_2RS_2)\\
&+ \log\mathcal{N}(\gamma_1; \mathbf{0}, \sigma_1^2\mathrm{I}) + \log\mathcal{N}(\gamma_3; \mathbf{0}, \sigma_3^2\mathrm{I})\\
\propto& -\frac{1}{2(1-\rho_1^2)}(\gamma_1^\top S_1^{-1}RS_1^{-1}\gamma_1 - 2\gamma_1^\top S_1^{-1}S_1^{-1}\hat{\gamma}_1 + 2\rho_1(\Gamma_1^\top+\beta_4\gamma_3^\top)S_3^{-1}S_1^{-1}\hat{\gamma}_1 + 2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\gamma_1)\\
&- \frac{1}{2(1-\rho_1^2)}(-2\rho(\Gamma_1^\top+\beta_4\gamma_3^\top)S_3^{-1}RS_1^{-1}\gamma_1 - 2(\Gamma_1^\top+\beta_4\gamma_3^\top)S_3^{-1}S_3^{-1}\hat{\Gamma}_1 + (\Gamma_1^\top+\beta_4\gamma_3^\top)S_3^{-1}RS_3^{-1}(\Gamma_1+\beta_4\gamma_3))\\
&- \frac{1}{2}((L_2^\top)^{-1}\hat{\gamma}_3 - U_2\gamma_3)^\top((L_2^\top)^{-1}\hat{\gamma}_3 - U_2\gamma_3) - \frac{\gamma_{3j}^2}{2\sigma_3^2}\\
\propto& -\frac{1}{2(1-\rho_1^2)}(\beta_4^2\gamma_3^\top S_3^{-1}RS_3^{-1}\gamma_3)\\
&- \frac{1}{2(1-\rho_1^2)}(2\rho_1\beta_4\gamma_3^\top S_3^{-1}S_1^{-1}\hat{\gamma}_1 - 2\beta_4\gamma_3^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1)\\
&- \frac{1}{2(1-\rho_1^2)}(-2\rho_1\beta_4\gamma_3^\top S_3^{-1}RS_1^{-1}\gamma_1 + 2\beta_1\beta_4\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_3 + 2\beta_4\boldsymbol{\beta}_2^\top S_3^{-1}RS_3^{-1}\gamma_3)\\
&- \frac{1}{2}((L_2^\top)^{-1}\hat{\gamma}_3 - U_2\gamma_3)^\top((L_2^\top)^{-1}\hat{\gamma}_3 - U_2\gamma_3) - \frac{\gamma_{3j}^2}{2\sigma_3^2}\\
=& -\frac{1}{2}\Big(\beta_4^2\frac{(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)} + \frac{1}{\sigma_3^2} + \sum_{j'}(U_2)_{j'j}^2\Big)\gamma_{3j}^2\\
&- \beta_4^2\frac{\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\gamma_{3j'}\gamma_{3j}}{(1-\rho_1^2)}\\
&- \frac{\rho_1\beta_4\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)} + \frac{\beta_4\sum_{j'}\hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)}\\
&+ \frac{\rho_1\beta_4\sum_{j'}\gamma_{1j'}(S_1^{-1}RS_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)} - \frac{\beta_1\beta_4\sum_{j'}\gamma_{1j'}(S_3^{-1}RS_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)} - \frac{\beta_4\sum_{j'}\beta_{2j'}(S_3^{-1}RS_3^{-1})_{j'j}\gamma_{3j}}{(1-\rho_1^2)}\\
&+ \Big(\sum_{j'}\tilde{\gamma}_{3j'}(U_2)_{j'j}\Big)\gamma_{3j} - \sum_{j'}\sum_{j''\neq j}(U_2)_{j'j}(U_2)_{j'j''}\gamma_{3j''}\gamma_{3j}.
\end{aligned}
$$

The conditional posterior distribution is $\gamma_{3j} \mid \Omega \backslash \gamma_{3j} \sim \mathcal{N}(\mu_{3j}, \sigma_{3j}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{3j}^2} =& -\frac{1}{2}(\beta_4^2 \frac{(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)} + \frac{1}{\sigma_3^2} + \sum_{j'}(U_2)_{j'j}^2), \\
\frac{\mu_{3j}}{\sigma_{3j}^2} =& -\frac{\sum_{j' \neq j}(S_3^{-1}RS_3^{-1})_{jj'}\gamma_{3j'}}{(1-\rho_1^2)} \\
& - \frac{\rho_1\beta_4 \sum_{j'} \hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} + \frac{\beta_4 \sum_{j'} \hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
& + \frac{\rho_1\beta_4 \sum_{j'} \gamma_{1j'}(S_1^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_1\beta_4 \sum_{j'} \gamma_{1j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_4 \sum_{j'} \beta_{2j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} \\
& + (\sum_{j'} \tilde{\gamma}_{3j'}(U_2)_{j'j} - \sum_{j'}\sum_{j'' \neq j}(U_2)_{j'j}(U_2)_{j'j''}\gamma_{3j''}.
\end{aligned}
$$

The full conditional distribution of β_1 is

$$
\begin{aligned}
& \log \mathcal{N}(\begin{pmatrix} \hat{\gamma}_1 \\ \hat{\Gamma}_1 \end{pmatrix}; \begin{pmatrix} S_1RS_1^{-1}\gamma_1 \\ S_3RS_3^{-1}(\Gamma_1 + \beta_4\gamma_3) \end{pmatrix}, \begin{pmatrix} S_1RS_1 & \rho_1 S_1RS_3 \\ \rho_1 S_1RS_3 & S_3RS_3 \end{pmatrix}) + \log \Pr(\beta_1) \\
\propto & -\frac{1}{2(1-\rho_1^2)}(\gamma_1^\top S_1^{-1}RS_1^{-1}\gamma_1 - 2\gamma_1^\top S_1^{-1}S_1^{-1}\hat{\gamma}_1 + 2\rho_1(\Gamma_1^\top + \beta_4\gamma_3^\top)S_3^{-1}S_1^{-1}\hat{\gamma}_1 + 2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\gamma_1) \\
& -\frac{1}{2(1-\rho_1^2)}(-2\rho(\Gamma_1^\top + \beta_4\gamma_3^\top)S_3^{-1}RS_1^{-1}\gamma_1 - 2(\Gamma_1^\top + \beta_4\gamma_3^\top)S_3^{-1}S_3^{-1}\hat{\Gamma}_1 + (\Gamma_1^\top + \beta_4\gamma_3^\top)S_3^{-1}RS_3^{-1}(\Gamma_1 + \beta_4\gamma_3)) \\
\propto & -\frac{1}{2}(\frac{\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_1}{(1-\rho_1^2)})\beta_1^2 \\
& -\frac{\rho_1\gamma_1^\top S_3^{-1}S_1^{-1}\hat{\gamma}_1 - \gamma_1^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1}{(1-\rho_1^2)}\beta_1 \\
& -\frac{-\rho_1\gamma_1^\top S_3^{-1}RS_1^{-1}\gamma_1 + \gamma_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\beta}_2 + \beta_4\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_3}{(1-\rho_1^2)}\beta_1.
\end{aligned}
$$

The conditional posterior distribution is $\beta_1 \mid \Omega \backslash \beta_1 \sim \mathcal{N}(\mu_{b1}, \sigma_{b1}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{b1}^2} =& -\frac{1}{2}(\frac{\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_1}{(1-\rho_1^2)}), \\
\frac{\mu_{b1}}{\sigma_{b1}^2} =& -\frac{\rho_1\gamma_1^\top S_3^{-1}S_1^{-1}\hat{\gamma}_1 - \gamma_1^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1}{(1-\rho_1^2)} \\
& -\frac{-\rho_1\gamma_1^\top S_3^{-1}RS_1^{-1}\gamma_1 + \gamma_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\beta}_2 + \beta_4\gamma_1^\top S_3^{-1}RS_3^{-1}\gamma_3}{(1-\rho_1^2)}.
\end{aligned}
$$

The full conditional distribution of β_{2j} is

$$
\begin{aligned}
&\log \mathcal{N}(\begin{pmatrix}\hat{\boldsymbol{\gamma}}_1\\ \hat{\Gamma}_1\end{pmatrix}; \begin{pmatrix}S_1 R S_1^{-1}\boldsymbol{\gamma}_1\\ S_3 R S_3^{-1}(\Gamma_1+\beta_4\boldsymbol{\gamma}_3)\end{pmatrix}, \begin{pmatrix}S_1 R S_1 & \rho_1 S_1 R S_3\\ \rho_1 S_1 R S_3 & S_3 R S_3\end{pmatrix}) + \log \mathcal{N}(\boldsymbol{\beta}_2; \mathbf{0}, \sigma_2^2 \mathrm{I})\\
\propto & -\frac{1}{2(1-\rho_1^2)}(\boldsymbol{\gamma}_1^\top S_1^{-1} R S_1^{-1}\boldsymbol{\gamma}_1 - 2\boldsymbol{\gamma}_1^\top S_1^{-1} S_1^{-1}\hat{\boldsymbol{\gamma}}_1 + 2\rho_1(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1 + 2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\boldsymbol{\gamma}_1)\\
& -\frac{1}{2(1-\rho_1^2)}(-2\rho(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1 - 2(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}S_3^{-1}\hat{\Gamma}_1 + (\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}RS_3^{-1}(\Gamma_1+\beta_4\boldsymbol{\gamma}_3))\\
& -\frac{\beta_{2j}^2}{2\sigma_2^2}\\
\propto & -\frac{1}{2(1-\rho_1^2)}(\boldsymbol{\beta}_2^\top S_3^{-1}RS_3^{-1}\boldsymbol{\beta}_2)\\
& -\frac{1}{2(1-\rho_1^2)}(2\rho_1\boldsymbol{\beta}_2^\top S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1 - 2\boldsymbol{\beta}_2^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1)\\
& -\frac{1}{2(1-\rho_1^2)}(-2\rho_1\boldsymbol{\gamma}_1^\top S_3^{-1}RS_1^{-1}\boldsymbol{\beta}_2 + 2\beta_1\boldsymbol{\gamma}_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\beta}_2 + 2\beta_4\boldsymbol{\beta}_2^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3)\\
& -\frac{\beta_{2j}^2}{2\sigma_2^2}\\
= & -\frac{1}{2}(\frac{(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)}+\frac{1}{\sigma_2^2})\beta_{2j}^2\\
& -\frac{\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\beta_{2j'}\beta_{2j}}{(1-\rho_1^2)}\\
& -\frac{\rho_1\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}\beta_{2j}}{(1-\rho_1^2)} + \frac{\sum_{j'}\hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}\beta_{2j}}{(1-\rho_1^2)}\\
& +\frac{\rho_1\sum_{j'}\gamma_{1j'}(S_1^{-1}RS_3^{-1})_{j'j}\beta_{2j}}{(1-\rho_1^2)} - \frac{\beta_1\sum_{j'}\gamma_{1j'}(S_3^{-1}RS_3^{-1})_{j'j}\beta_{2j}}{(1-\rho_1^2)} - \frac{\beta_4\sum_{j'}\gamma_{3j'}(S_3^{-1}RS_3^{-1})_{j'j}\beta_{2j}}{(1-\rho_1^2)}.
\end{aligned}
$$

The conditional posterior distribution is, $\beta_{2j} \mid \Omega\backslash\beta_{2j} \sim \mathcal{N}(\mu_{b2}, \sigma_{b2}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{b2}^2} =& -\frac{1}{2}(\frac{(S_3^{-1}RS_3^{-1})_{jj}}{(1-\rho_1^2)}+\frac{1}{\sigma_2^2}),\\
\frac{\mu_{b2}}{\sigma_{b2}^2} =& -\frac{\sum_{j'\neq j}(S_3^{-1}RS_3^{-1})_{jj'}\beta_{2j'}}{(1-\rho_1^2)}\\
& -\frac{\rho_1\sum_{j'}\hat{\gamma}_{1j'}(S_1^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)} + \frac{\sum_{j'}\hat{\Gamma}_{1j'}(S_3^{-1}S_3^{-1})_{j'j}}{(1-\rho_1^2)}\\
& +\frac{\rho_1\sum_{j'}\gamma_{1j'}(S_1^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_1\sum_{j'}\gamma_{1j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)} - \frac{\beta_4\sum_{j'}\gamma_{3j'}(S_3^{-1}RS_3^{-1})_{j'j}}{(1-\rho_1^2)}.
\end{aligned}
$$

The full conditional distribution of β_4 is

$$
\begin{aligned}
&\log\mathcal{N}(\begin{pmatrix}\hat{\boldsymbol{\gamma}}_1\\ \hat{\Gamma}_1\end{pmatrix};\begin{pmatrix}S_1RS_1^{-1}\boldsymbol{\gamma}_1\\ S_3RS_3^{-1}(\Gamma_1+\beta_4\boldsymbol{\gamma}_3)\end{pmatrix},\begin{pmatrix}S_1RS_1 & \rho_1S_1RS_3\\ \rho_1S_1RS_3 & S_3RS_3\end{pmatrix})+\log\Pr(\beta_4)\\
\propto&-\frac{1}{2(1-\rho_1^2)}(\boldsymbol{\gamma}_1^\top S_1^{-1}RS_1^{-1}\boldsymbol{\gamma}_1-2\boldsymbol{\gamma}_1^\top S_1^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1+2\rho_1(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1+2\rho_1\hat{\Gamma}_1^\top S_3^{-1}S_1^{-1}\boldsymbol{\gamma}_1)\\
&-\frac{1}{2(1-\rho_1^2)}(-2\rho(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1-2(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}S_3^{-1}\hat{\Gamma}_1+(\Gamma_1^\top+\beta_4\boldsymbol{\gamma}_3^\top)S_3^{-1}RS_3^{-1}(\Gamma_1+\beta_4\boldsymbol{\gamma}_3))\\
\propto&-\frac{1}{2}(\frac{\boldsymbol{\gamma}_3^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3}{(1-\rho_1^2)})\beta_4^2\\
&-\frac{\rho_1\boldsymbol{\gamma}_3^\top S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1-\boldsymbol{\gamma}_3^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1}{(1-\rho_1^2)}\beta_4\\
&-\frac{-\rho_1\boldsymbol{\gamma}_3^\top S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1+\beta_1\boldsymbol{\gamma}_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3+\boldsymbol{\beta}_2^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3}{(1-\rho_1^2)}\beta_4.
\end{aligned}
$$

The conditional posterior distribution is, $\beta_4 \mid \Omega\backslash\beta_4 \sim \mathcal{N}(\mu_{b4},\sigma_{b4}^2)$, where

$$
\begin{aligned}
-\frac{1}{2\sigma_{b4}^2}=&-\frac{1}{2}(\frac{\boldsymbol{\gamma}_3^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3}{(1-\rho_1^2)}),\\
\frac{\mu_{b4}}{\sigma_{b4}^2}=&-\frac{\rho_1\boldsymbol{\gamma}_3^\top S_3^{-1}S_1^{-1}\hat{\boldsymbol{\gamma}}_1-\boldsymbol{\gamma}_3^\top S_3^{-1}S_3^{-1}\hat{\Gamma}_1}{(1-\rho_1^2)}\\
&-\frac{-\rho_1\boldsymbol{\gamma}_3^\top S_3^{-1}RS_1^{-1}\boldsymbol{\gamma}_1+\beta_1\boldsymbol{\gamma}_1^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3+\boldsymbol{\beta}_2^\top S_3^{-1}RS_3^{-1}\boldsymbol{\gamma}_3}{(1-\rho_1^2)}.
\end{aligned}
$$

The posterior distribution of $\sigma_1^2,\sigma_2^2,\sigma_3^2$ is the same as MERLIN above.

S5 Modeling direct SNP×E effects in the outcome

We extend the outcome model to incorporate direct SNP×E effects:

$$
Y_i = X_i\beta^{(A)}+\sum_j G_{ij}\beta_j^{(G)}+\sum_j G_{ij}E_i\beta_j^{(GI)}+E_i\beta^{(E)}+X_iE_i\beta^{(I)}+U_i\eta_Y+\epsilon_{Y_i},
$$

where $\beta_j^{(GI)}$ represents the horizontal pleiotropic interaction effect. Consequently, the total SNP×E effect on Y (i.e., the parameter estimated by an outcome GWIS), denoted as Γ_{3*}, is given by:

$$
\Gamma_{3*}=\beta^{(A)}\boldsymbol{\gamma}^{(GI)}+\beta^{(I)}\boldsymbol{\gamma}^{(G)}+\boldsymbol{\beta}^{(GI)}.
$$

Under this specification, the likelihood (8) of the observed data can be written as follows:

$$
\begin{aligned}
L(\theta|\mathcal{D})=&\log\mathcal{N}(\begin{pmatrix}\hat{\boldsymbol{\gamma}}_1\\ \hat{\Gamma}_1\end{pmatrix};\begin{pmatrix}S_1RS_1^{-1}\boldsymbol{\gamma}_1\\ S_3RS_3^{-1}\Gamma_1+S_3^2S_4^{-1}RS_4^{-1}\beta_4\boldsymbol{\gamma}_3\end{pmatrix},\begin{pmatrix}S_1RS_1 & \rho_1S_1RS_3\\ \rho_1S_1RS_3 & S_3RS_3\end{pmatrix})\\
&+\log\mathcal{N}(\begin{pmatrix}\hat{\boldsymbol{\gamma}}_3\\ \hat{\Gamma}_3\end{pmatrix};\begin{pmatrix}S_2RS_2^{-1}\boldsymbol{\gamma}_3\\ S_4RS_4^{-1}\Gamma_{3*}\end{pmatrix},\begin{pmatrix}S_2RS_2 & \rho_2S_2RS_4\\ \rho_2S_2RS_4 & S_4RS_4\end{pmatrix})\\
&+\log\mathcal{N}(\boldsymbol{\gamma}_1;\mathbf{0},\sigma_1^2\mathrm{I})+\log\mathcal{N}(\boldsymbol{\beta}_2;0,\sigma_2^2\mathrm{I})+\log\mathcal{N}(\boldsymbol{\gamma}_3;\mathbf{0},\sigma_3^2\mathrm{I})+\log\mathcal{N}(\boldsymbol{\beta}^{(GI)};0,\sigma_4^2\mathrm{I}).
\end{aligned}
\tag{11}
$$

Analogous to $\boldsymbol{\gamma}_1,\boldsymbol{\beta}_2$, and $\boldsymbol{\gamma}_3$, we assume $\boldsymbol{\beta}^{(GI)}$ follows a normal prior with mean 0 and variance $\sigma_4^2\mathrm{I}$. These derivations extend naturally to scenarios involving sample overlap or continuous environmental variables.

S6 MERLIN for binary traits

Here, we show that the MERLIN framework is applicable to binary traits in case–control studies. Following Hu et al. (2022)[5], we detail the scenarios that involve either a binary exposure or a binary outcome.

We begin with the linear structural full model for the continuous exposure x_i and outcome y_i for individual i:

$$x_i = \beta_{0,x} + \boldsymbol{C}_i^\top \beta_{cov,x} + \sum_j G_{ij}\gamma_j^{(G)} + \sum_j G_{ij}E_i\gamma_j^{(GI)} + E_i\gamma^{(E)}, + U_i\eta_x + \epsilon_{x_i}, \tag{12}$$

$$y_i = \beta_{0,y} + \boldsymbol{C}_i^\top \beta_{cov,y} + x_i\beta^{(A)} + \sum_j G_{ij}\beta_j^{(G)} + E_i\beta^{(E)} + x_iE_i\beta^{(I)} + U_i\eta_y + \epsilon_{y_i}, \tag{13}$$

where $\beta_{0,x}$ and $\beta_{0,y}$ denote the intercept, $\boldsymbol{C}_i$ represents the covariate vector with corresponding effects $\beta_{cov,x}$ and $\beta_{cov,y}$. G_i denotes a centered genotype, E_i is the centered environmental variable, U_i is the unmeasured confounder common to both $\boldsymbol{x}$ and $\boldsymbol{y}$, and is assumed to be independent of G_i and E_i, $\boldsymbol{\gamma}^{(G)}$ is the average effect of SNPs on exposure $\boldsymbol{x}$, $\gamma^{(E)}$ is the average effect of environment E on exposure $\boldsymbol{x}$, $\gamma^{(GI)}$ is the GE interaction effect of SNPs on exposure $\boldsymbol{x}$, $\beta^{(A)}$ is the average causal interaction effect of $\boldsymbol{x}$ on $\boldsymbol{y}$, $\beta^{(G)}$ is the horizontal pleiotropic effect of SNPs on outcome $\boldsymbol{y}$, $\beta^{(E)}$ is the average effect of modifier E on outcome $\boldsymbol{y}$, $\beta^{(I)}$ is the causal interaction effect of $\boldsymbol{x} \times E$ on $\boldsymbol{y}$, representing how the effect of $\boldsymbol{x}$ on $\boldsymbol{y}$ is modified by E, η_x and η_y are the effects of U_i on $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively, ϵ_x and ϵ_y are independent random errors and are assumed independent of other terms in their respective equations.

Case 1: a continuous exposure ($\boldsymbol{x}$) and a binary outcome ($\boldsymbol{y}$).

For binary outcome traits $\boldsymbol{y}$, we assume the following probit model and insert (12) into it [6]:

$$p(y_i = 1 \mid G_{ij}, E_i, \boldsymbol{C}_i) \tag{14}$$

$$= \Phi(\beta_{0,y}^b + \boldsymbol{C}_i^\top \beta_{cov,y}^b + x_i\beta^{(A),b} + \sum_j G_{ij}\beta_j^{(G),b} + E_i\beta^{(E),b} + x_iE_i\beta^{(I),b} + U_i\eta_y^b + \epsilon_{y_i}) \tag{15}$$

$$= \Phi(\beta_0^b + \boldsymbol{C}_i^\top \beta_{cov}^b + \sum_j G_{ij}(\beta^{(A),b}\gamma_j^{(G)} + \beta_j^{(G),b}) + \sum_j G_{ij}E_i(\beta^{(A),b}\gamma_j^{(GI)} + \beta^{(I),b}\gamma_j^{(G)}) + \sum_j G_{ij}E_i^2\beta_{GE^2}^b + E_i\beta_E^b + E_i^2\beta_{E^2}^b + \varepsilon_{y_i}), \tag{16}$$

$$= \Phi(\beta_0^b + \boldsymbol{C}_i^\top \beta_{cov}^b + \sum_j G_{ij}\Gamma_j^{(G),b} + \sum_j G_{ij}E_i\Gamma_j^{(GI),b} + \sum_j G_{ij}E_i^2\beta_{GE^2}^b + E_i\beta_E^b + E_i^2\beta_{E^2}^b + \varepsilon_{y_i}), \tag{17}$$

where $\Phi()$ is the cumulative distribution function of the standard normal distribution. In equation (15), represents the probit model for phenotype $\boldsymbol{y}$. Each term in $\Phi()$ has a meaning similar to that in Equation (13), but the superscript 'b' denotes that these are the true causal and genetic effects on the liability scale. Substituting x_i into (15) yields equation (16), which simplifies to equation (17), where $\beta_0^b = \beta_{0,y}^b + \beta^{(A),b}\beta_{0,x}$, and $\beta_{cov}^b = \beta_{cov,y}^b + \beta^{(A),b}\beta_{cov,x}$. The parameters $\beta_{GE^2}^b, \beta_E^b, \beta_{E^2}^b$ are the effect sizes of the corresponding variables. It is noteworthy that the parameters $\beta^{(E),b}$ and β_E^b capture conceptually different effects. The parameter $\beta^{(E),b}$ quantifies the direct association between the environment variable E and the outcome y in the structural model for y, whereas β_E^b represents the resulting coefficient of E in the outcome model after substituting the structural equation of x into that of y and algebraically simplifying the expression. The error term $\varepsilon_{y_i} = E_iU_i\eta_x\beta^{(I)} + U_i(\eta_x\beta^{(A)} + \eta_y) + \epsilon_x\beta^{(A)} + E_i\epsilon_x\beta^{(I)} + \epsilon_y$. For clarity, let $\Gamma_j^{(G),b} = \beta^{(A),b}\gamma_j^{(G)} + \beta_j^{(G),b}$, and $\Gamma_j^{(GI),b} = \beta^{(A),b}\gamma_j^{(GI)} + \beta^{(I),b}\gamma_j^{(G)}$.

Applying the known results in [6, 7], we have a linear approximation of $p(y_i = 1 \mid G_{ij}, E_i)$ as

$$p(y_i = 1 \mid G_i, E_i) \approx k + \frac{k(1-k)\phi(\beta_0^b)}{K(1-K)}(\boldsymbol{C}_i^\top \beta_{cov}^b + \sum_j G_{ij}\Gamma_j^{(G),b} + \sum_j G_{ij}E_i\Gamma_j^{(GI),b} + \sum_j G_{ij}E_i^2\beta_{GE^2}^b + E_i\beta_E^b + E_i^2\beta_{E^2}^b + \varepsilon_{y_i}),$$

where k and K represent the proportions of the cases in the ascertained case-control sample and the population, respectively. This implies that the effect sizes estimated by the linear model, $\Gamma_j^{(G)}$ and

$\Gamma_j^{(GI)}$, can be transformed into the liability scale by

$$\Gamma_j^{(G),b} = \frac{K(1-K)}{k(1-k)\phi(\beta_0^b)}\Gamma_j^{(G)} \text{ and } \Gamma_j^{(GI),b} = \frac{K(1-K)}{k(1-k)\phi(\beta_0^b)}\Gamma_j^{(GI)}.$$

Consequently, we have

$$\beta^{(A),b} = \frac{K(1-K)}{k(1-k)\phi(\beta_0^b)}\beta^{(A)} \text{ and } \beta^{(I),b} = \frac{K(1-K)}{k(1-k)\phi(\beta_0^b)}\beta^{(I)}.$$

This direct proportionality means that a hypothesis test for the null effect on the observed scale (e.g. $H_0 : \beta^{(I)} = 0$) is equivalent to a test for the null effect on the liability scale ($H_0 : \beta^{(I),b} = 0$). Therefore, the significance tests for the main and heterogeneity effects produced by MERLIN can be directly interpreted as tests for the existence of these causal effects on the underlying liability to the phenotype.

Case 2: a binary exposure and a continuous outcome.

For a binary exposure $\boldsymbol{x}$, we again consider the following probit model:

$$p(x_i = 1 \mid G_{ij}, E_i) = \Phi(\beta_{0,x}^b + \boldsymbol{C}_i^\top \beta_{cov,x}^b + \sum_j G_{ij}\gamma_j^{(G),b} + \sum_j G_{ij}E_i\gamma_j^{(GI),b} + E_i\gamma^{(E),b} + U_i\eta_x^b + \epsilon_{x_i}),$$

where each term in $\Phi()$ corresponds to the same interpretation as the phenotype x_i described in equation (12), but add the superscript 'b' to indicate that these are the true causal and genetic effects on the liability scale.

With the above preparation, we can apply the known results in [6, 7] to obtain an approximation of $p(x_i = 1 \mid G_{ij}, E_i)$ as

$$p(x_i = 1 \mid G_{ij}, E_i) \approx k + \frac{k(1-k)\phi(\beta_0^b)}{K(1-K)}(\boldsymbol{C}_i^\top \beta_{cov,x}^b + \sum_j G_{ij}\gamma_j^{(G),b} + \sum_j G_{ij}E_i\gamma_j^{(GI),b} + E_i\gamma^{(E),b} + \varepsilon_{x_i}), \tag{18}$$

where k and K represent the proportions of the cases in the ascertained case-control sample and the population, respectively. $\varepsilon_{x_i} = U_i\eta_x^b + \epsilon_{x_i}$. Equation (18) implies that the estimates of $\gamma_j^{(G)}$ and $\gamma_j^{(GI)}$ using a linear model can be transformed into the liability scale by

$$\gamma_j^{(G),b} = \frac{K(1-K)}{k(1-k)\phi(\beta_0^b)}\gamma_j^{(G)} \text{ and } \gamma_j^{(GI),b} = \frac{K(1-K)}{k(1-k)\phi(\beta_0^b)}\gamma_j^{(GI)}.$$

Plugging (18) into a continuous y (13), we have

$$y_i \approx \beta_0 + \boldsymbol{C}_i^\top \beta_{cov} + \sum_j G_{ij}\Gamma_j^{(G),b} + \sum_j G_{ij}E_i\Gamma_j^{(GI),b} + \sum_j G_{ij}E_i^2\beta_{GE^2} + E_i\beta_E + E_i^2\beta_{E^2} + \varepsilon_{y_i},$$

where $\beta_0 = \beta_{0,y} + k$, $\beta_{cov} = \beta_{cov,y} + \frac{k(1-k)\phi(\beta_0)}{K(1-K)}\beta^{(A)}\beta_{cov,x}$, $\Gamma_j^{(G),b} = \beta^{(A)}\gamma_j^{(G),b} + \beta_j^{(G)}$, $\Gamma_j^{(GI),b} = \beta^{(A)}\gamma_j^{(GI),b} + \beta^{(I)}\gamma_j^{(G),b}$, $\varepsilon_{y_i} = \frac{k(1-k)\phi(\beta_0^b)}{K(1-K)} * E_i\beta^{(I)}\varepsilon_{x_i} + \frac{k(1-k)\phi(\beta_0^b)}{K(1-K)} * \beta^{(A)}\varepsilon_{x_i} + U_i\eta_y + \epsilon_y$.The other parameters carry meanings analogous to those in (13).

Case 3: a binary exposure and a binary outcome.

Similarly, for binary exposure x and outcome y, we consider the probit model as follows:

$$\begin{aligned} p(x_i = 1 \mid G_{ij}, E_i) &= \Phi(\beta_{0,x}^b + \boldsymbol{C}_i^\top \beta_{cov,x}^b + \sum_j G_{ij}\gamma_j^{(G),b} + \sum_j G_{ij}E_i\gamma_j^{(GI),b} + E_i\gamma^{(E),b} + \varepsilon_{x_i}) \\ &\approx k_x + \frac{k_x(1-k_x)\phi(\beta_0^b)}{K_x(1-K_x)}(\boldsymbol{C}_i^\top \beta_{cov,x}^b + \sum_j G_{ij}\gamma_j^{(G),b} + \sum_j G_{ij}E_i\gamma_j^{(GI),b} + E_i\gamma^{(E),b} + \varepsilon_{x_i}), \end{aligned} \tag{19}$$

$$p(y_i = 1 \mid G_i, E_i) = \Phi(\beta_{0,y}^b + \boldsymbol{C}_i^\top \beta_{cov,y}^b + x_i\beta^{(A),b} + \sum_j G_{ij}\beta_j^{(G),b} + E_i\beta^{(E),b} + x_iE_i\beta^{(I),b} + U_i\eta_y^b + \epsilon_{y_i}), \tag{20}$$

Equations (19) and Equations (18), Equations (20) and Equations (15) are mathematically equivalent, respectively. However, to distinguish the notation, we represent k as k_x and k_y, respectively. Plugging (19) into (20), we obtain the following.

$$p(y_i = 1 \mid G_i, E_i) = \Phi(\beta_0^b + \boldsymbol{C}_i^\top \beta_{cov}^b + \sum_j G_{ij}\Gamma_j^{(G),b} + \sum_j G_{ij}E_i\Gamma_j^{(GI),b} + \sum_j G_{ij}E_i^2\beta_{GE^2}^b + E_i\beta_E^b + E_i^2\beta_{E^2}^b + \varepsilon_{y_i}),$$

where $\beta_0^b = \beta_{0,y}^b + \beta^{(A),b}k_x$, $\beta_{cov}^b = \beta_{cov,y}^b + \frac{k_x(1-k_x)\phi(\beta_0^b)}{K_x(1-K_x)}\beta^{(A),b}\beta_{cov,x}^b$, $\Gamma_j^{(G),b} = \frac{k_x(1-k_x)\phi(\beta_0^b)}{K_x(1-K_x)}\beta^{(A),b}\gamma_j^{(G),b} + \beta_j^{(G),b}$, $\Gamma_j^{(GI),b} = \frac{k_x(1-k_x)\phi(\beta_0^b)}{K_x(1-K_x)}(\beta^{(A),b}\gamma_j^{(GI),b} + \beta^{(I),b}\gamma_j^{(G),b})$, $\beta_{GE^2}^b, \beta_E^b, \beta_{E^2}^b$ are the coefficients of the variables that correspond to each other. $\varepsilon_{y_i} = \frac{k_x(1-k_x)\phi(\beta_0^b)}{K_x(1-K_x)} * E_i\beta^{(I),b}\varepsilon_{x_i} + \frac{k_x(1-k_x)\phi(\beta_0^b)}{K_x(1-K_x)} * \beta^{(A),b}\varepsilon_{x_i} + U_i\eta_y + \epsilon_y$. A similar approach to the previous one, we can obtain

$$\Gamma_j^{(G),b} = \frac{K_y(1-K_y)}{k_y(1-k_y)\phi(\beta_0^b)}\Gamma_j^{(G)} \text{ and } \Gamma_j^{(GI),b} = \frac{K_y(1-K_y)}{k_y(1-k_y)\phi(\beta_0^b)}\Gamma_j^{(GI)},$$

$$\beta^{(A),b} = \frac{K_y(1-K_y)}{k_y(1-k_y)\phi(\beta_0^b)}\beta^{(A)} \text{ and } \beta^{(I),b} = \frac{K_y(1-K_y)}{k_y(1-k_y)\phi(\beta_0^b)}\beta^{(I)}.$$

Based on the derivation above, the causal effect estimated using linear models remains meaningful and interpretable in all three scenarios.

S7 Evaluating bias in standard MR average effect estimates under real-data heterogeneity

Our primary simulations (Results Section 2.2) demonstrate that standard MR methods can produce biased estimates of the average causal effect in the presence of causal heterogeneity ($\beta_4 \neq 0$). To demonstrate this phenomenon in a real-world context, we examined the SCZ→IDP.0664 assocation, which our main analysis identified as having a significant sex-specific effects.

A challenge in real data is that the true average effect (β_1) is unknown. We therefore derived a more robust proxy for the true average effect by first performing sex-stratified MR analyses to obtain male-specific ($\hat{\beta}_{\mathrm{M}}$) and female-specific ($\hat{\beta}_{\mathrm{F}}$) causal estimates, and then averaging them: $\hat{\beta}_1 = 0.5(\hat{\beta}_{\mathrm{M}} + \hat{\beta}_{\mathrm{F}})$. This approach is expected to be less biased than a naive combined-sex analysis because it explicitly accounts for the identified heterogeneity.

We then compared the estimates from standard MR (applied directly to combined-sex summary statistics) with this sex-stratified combined estimate (Supplementary Fig. S15a). Across all four comparator methods (IVW, MR-Egger, RAPS, MR-LDP), the standard MR estimates (green lines) were systematically different from the corresponding sex-stratified combined estimates (blue lines).

To verify that this observed divergence represents bias in the standard MR approach, we conducted a targeted simulation study using parameters informed by the real-data analysis (setting the true β_1 and β_4 to the values obtained from MERLIN). The results (Supplementary Fig. S15b) confirm this hypothesis. The boxplots for the standard MR methods are clearly biased (shifted away from the true value indicated by the dashed line), while the boxplots for the sex-stratified combined MR approach are closer to the true average effect.

This real-data example, supported by targeted simulation, therefore provides strong evidence that applying standard MR methods to combined data in the presence of significant sex heterogeneity can lead to biased estimation of the average effect. This highlights the critical importance of using methods that can appropriately model such heterogeneity effects.

S8 Supplementary Discussion: Biological Context of Causal Findings in Section 2.6

Here, we provide a more detailed interpretation of our significant MR findings, connecting them with existing neurobiological literature.

S8.1 Causal Effect of Cerebellar Volume on Schizophrenia Risk

Our forward MR finding, indicating that increased volume of the left cerebellar cortex (IDP.0194) elevates SCZ risk, aligns with the cerebellum's recognized role in cognitive and emotional processes disrupted in SCZ. This involvement is highlighted by functional neuroimaging studies [8, 9] and conceptualized within frameworks such as cognitive dysmetria, which implicates cortico-cerebellar-thalamo-cortical (CCTC) circuits in the pathophysiology of the disorder [10]. While reports on cerebellar volume changes in SCZ have been varied across studies, with some finding decreases and others increases or no change, our specific causal finding points to a potentially pathogenic role for volumetric increases in this specific cerebellar subregion, warranting further investigation into circuit-level disruptions.

S8.2 Causal Effects of Schizophrenia Liability on Brain Structure

Conversely, our analyses show that SCZ liability exerts causal effects on multiple sensorimotor domains. The finding of reduced cortical surface area in the left paracentral lobule (IDP.0664) aligns with evidence suggesting early neurodevelopmental deficits in somatomotor integration in individuals with SCZ [11]. Furthermore, we observed microstructural alterations in the cerebral peduncles, marked by a decreased orientation dispersion index (IDP.1991, IDP.1992) and an elevated mode of anisotropy (IDP.1541, right). The cerebral peduncles contain major corticofugal tracts (e.g., corticospinal, corticobulbar). These findings likely reflect white matter reorganization within these pathways, consistent with diffusion tensor imaging studies reporting widespread white matter alterations in both first-episode [12] and treatment-resistant [13] SCZ, which may also involve hemispheric asymmetries [14]. Such diffusion metric alterations may point towards progressive axonal changes (e.g., changes in density or myelination) and could underlie both motor coordination deficits (via disrupted efferent signaling) and cognitive disorganization (via impaired cortico–subcortical communication) [12, 14].

S8.3 Sex-Specific Causal Effects of Schizophrenia Liability

Our most striking findings were the significant sex differences in the causal effects of SCZ on the surface area of the left paracentral lobule (IDP.0664) and the volume of the left nucleus accumbens (IDP.0015). These sex-specific patterns are highly plausible given the extensive literature documenting sex differences in brain structure, cognition, and clinical presentation in SCZ [15]. Our results suggest that schizophrenia may exert more detrimental effects on neural circuits relevant to cognition and behavior in males. The paracentral lobule is critical for sensorimotor function, and its alteration may relate to motor abnormalities often seen in SCZ. The nucleus accumbens is a core component of the brain's reward system, and its dysfunction is linked to negative symptoms and motivational deficits. The more pronounced volume reduction in males could contribute to observed sex differences in these clinical features or associated cognitive impairments [16]. This provides a potential causal neurobiological underpinning for why males often experience an earlier onset and more severe course of illness.

References

[1] Stephen Burgess, George Davey Smith, Neil M Davies, Frank Dudbridge, Dipender Gill, M Maria Glymour, Fernando P Hartwig, Zoltán Kutalik, Michael V Holmes, Cosetta Minelli, et al. Guidelines for performing mendelian randomization investigations: update for summer 2023. *Wellcome open research*, 4:186, 2023.

[2] Xiang Zhu and Matthew Stephens. Bayesian large-scale multiple regression with summary statistics from genome-wide association studies. *The annals of applied statistics*, 11(3):1561, 2017.

[3] Qing Cheng, Xiao Zhang, Lin S Chen, and Jin Liu. Mendelian randomization accounting for complex correlated horizontal pleiotropy while elucidating shared genetic etiology. *Nature communications*, 13(1):6490, 2022.

[4] Stefan Wilhelm and Maintainer Stefan Wilhelm. Package 'tmvtnorm'. *R journal*, 2015.

[5] Xianghong Hu, Jia Zhao, Zhixiang Lin, Yang Wang, Heng Peng, Hongyu Zhao, Xiang Wan, and Can Yang. Mendelian randomization for causal inference accounting for pleiotropy and sample structure using genome-wide summary statistics. *Proceedings of the National Academy of Sciences*, 119(28):e2106858119, 2022.

[6] Sang Hong Lee, Naomi R Wray, Michael E Goddard, and Peter M Visscher. Estimating missing heritability for disease from genome-wide association studies. *The American Journal of Human Genetics*, 88(3):294–305, 2011.

[7] Xiang Zhou, Peter Carbonetto, and Matthew Stephens. Polygenic modeling with bayesian sparse linear mixed models. *PLoS genetics*, 9(2):e1003264, 2013.

[8] X. Wang, Y. Zhang, J. Huang, Y. Wang, Y. Niu, S. S. Y. Lui, L. Hui, and R. C. K. Chan. Revisiting reward impairments in schizophrenia spectrum disorders: a systematic review and meta-analysis for neuroimaging findings. *Psychol Med*, 53(15):7189–7202, 2023. Wang, Xuan Zhang, Yinghao Huang, Jia Wang, Yi Niu, Yanzhe Lui, Simon S Y Hui, Li Chan, Raymond C K eng Meta-Analysis Systematic Review England 2023/03/31 Psychol Med. 2023 Nov;53(15):7189-7202. doi: 10.1017/S0033291723000703. Epub 2023 Mar 30.

[9] O. Lungu, M. Barakat, S. Laventure, K. Debas, S. Proulx, D. Luck, and E. Stip. The incidence and nature of cerebellar findings in schizophrenia: a quantitative review of fmri literature. *Schizophr Bull*, 39(4):797–806, 2013. Lungu, Ovidiu Barakat, Marc Laventure, Samuel Debas, Karen Proulx, Sebastien Luck, David Stip, Emmanuel eng Research Support, Non-U.S. Gov't Review 2012/01/24 Schizophr Bull. 2013 Jul;39(4):797-806. doi: 10.1093/schbul/sbr193. Epub 2012 Jan 20.

[10] N. C. Andreasen. A unitary model of schizophrenia -. *Archives of General Psychiatry*, 56(9):781–787, 1999. 232pk Times Cited:455 Cited References Count:96.

[11] S. Gao, Y. Ming, S. Ni, Z. Kong, J. Wang, Y. Gu, S. Lu, T. Chen, M. Kong, J. Sun, and X. Xu. Association of reduced local activities in the default mode and sensorimotor networks with clinical characteristics in first-diagnosed episode of schizophrenia. *Neuroscience*, 495:47–57, 2022. Gao, Shuzhan Ming, Yidan Ni, Sulin Kong, Zhiyao Wang, Jiayin Gu, Yuan Lu, Shuiping Chen, Tian Kong, Mingjun Sun, Jing Xu, Xijia eng Research Support, Non-U.S. Gov't 2022/05/24 Neuroscience. 2022 Jul 15;495:47-57. doi: 10.1016/j.neuroscience.2022.05.020. Epub 2022 May 20.

[12] A. V. Faria, J. Crawford, C. F. Ye, J. Hsu, A. Kenkare, D. Schretlen, and A. Sawa. Relationship between neuropsychological behavior and brain white matter in first-episode psychosis. *Schizophrenia Research*, 208:49–54, 2019. Hz7gq Times Cited:19 Cited References Count:57.

[13] R. Ochi, Y. Noda, S. Tsuchimoto, R. Tarumi, S. Honda, K. Matsushita, S. Tsugawa, E. Plitman, F. Masuda, K. Ogyu, M. Wada, T. Miyazaki, S. Fujii, M. M. Chakravarty, A. Graff-Guerrero, H. Uchida, M. Mimura, and S. Nakajima. White matter microstructural organizations in patients with severe treatment-resistant schizophrenia: A diffusion tensor imaging study. *Prog Neuropsychopharmacol Biol Psychiatry*, 100:109871, 2020. Ochi, Ryo Noda, Yoshihiro Tsuchimoto, Shohei Tarumi, Ryosuke Honda, Shiori Matsushita, Karin Tsugawa, Sakiko Plitman, Eric Masuda, Fumi Ogyu, Kamiyu Wada, Masataka Miyazaki, Takahiro Fujii, Shinya Chakravarty, M Mallar Graff-Guerrero, Ariel Uchida, Hiroyuki Mimura, Masaru Nakajima, Shinichiro eng Research Support, Non-U.S. Gov't England 2020/01/22 Prog Neuropsychopharmacol Biol Psychiatry. 2020 Jun 8;100:109871. doi: 10.1016/j.pnpbp.2020.109871. Epub 2020 Jan 18.

[14] N. F. Ho, Z. Li, F. Ji, M. Wang, C. N. Kuswanto, M. Y. Sum, H. Y. Tng, Y. Y. Sitoh, K. Sim, and J. Zhou. Hemispheric lateralization abnormalities of the white matter microstructure in patients with schizophrenia and bipolar disorder. *J Psychiatry Neurosci*, 42(4):242–251, 2017. Ho, New Fei Li, Zhengjun Ji, Fang Wang, Mingyuan Kuswanto, Carissa N Sum, Min Yi Tng, Han Ying Sitoh, Yih Yian Sim, Kang Zhou, Juan eng Research Support, Non-U.S. Gov't Canada 2017/02/25 J Psychiatry Neurosci. 2017 Jun;42(4):242-251. doi: 10.1503/jpn.160090.

[15] A. Mendrek and A. Mancini-Marie. Sex/gender differences in the brain and cognition in schizophrenia. *Neurosci Biobehav Rev*, 67:57–78, 2016. Mendrek, Adrianna Mancini-Marie, Adham eng Review 2016/01/09 Neurosci Biobehav Rev. 2016 Aug;67:57-78. doi: 10.1016/j.neubiorev.2015.10.013. Epub 2015 Dec 30.

[16] W. N. Zhang, D. M. Wang, H. X. Zhou, D. C. Chen, M. H. Xiu, L. X. Cui, and X. Y. Zhang. Sex difference in association between tardive dyskinesia and cognitive deficits in patients with chronic schizophrenia. *Schizophrenia Research*, 255:140–147, 2023. L9ha8 Times Cited:1 Cited References Count:52.

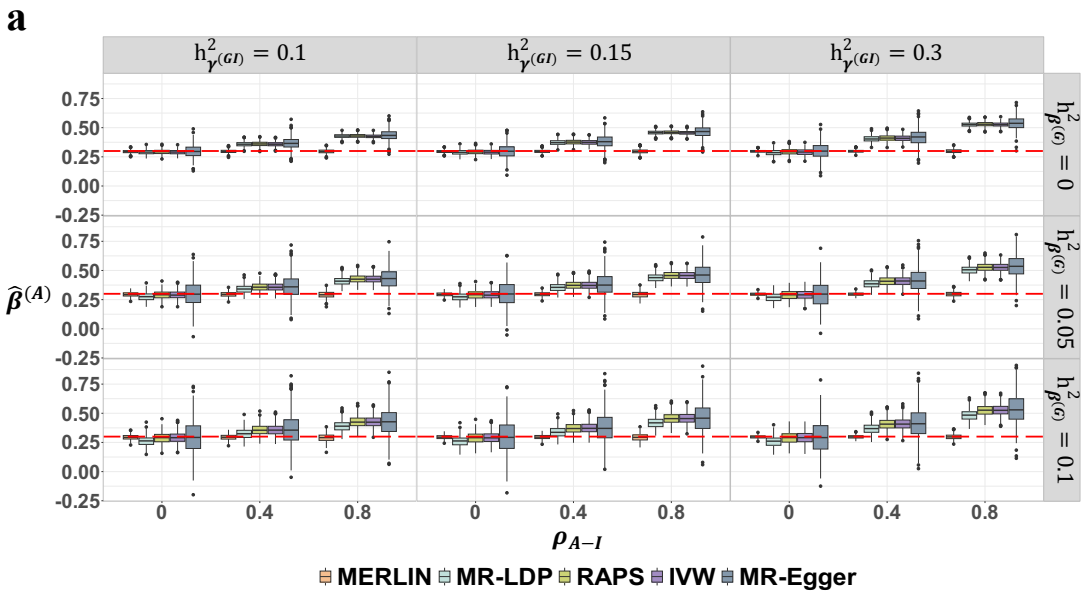

Figure S1. Simulation performance of MERLIN and standard MR methods for estimating the average causal effect ($\beta^{(A)}$). Simulation assumed a true average and interaction effect $\beta^{(A)} = \beta^{(I)} = 0.3$ and were based on 500 replicates. Standard MR methods included MR-LDP, RAPS, IVW, and MR-Egger. Boxplots of average effect estimates ($\beta^{(A)}$) are shown across varying proportions of exposure variance explained by G×E effects ($h^2_{\gamma(GI)} = 0.1, 0.15, 0.3$) and varying correlations between GWAS and GWIS instrument effects (ρ_{A-I}; x-axis within each plot: $0, 0.4, 0.8$), with horizontal pleiotropy ($h^2_{\beta(G)} = 0, 0.05, 0.1$). Dashed red lines indicate the true $\beta^{(A)}$ values.

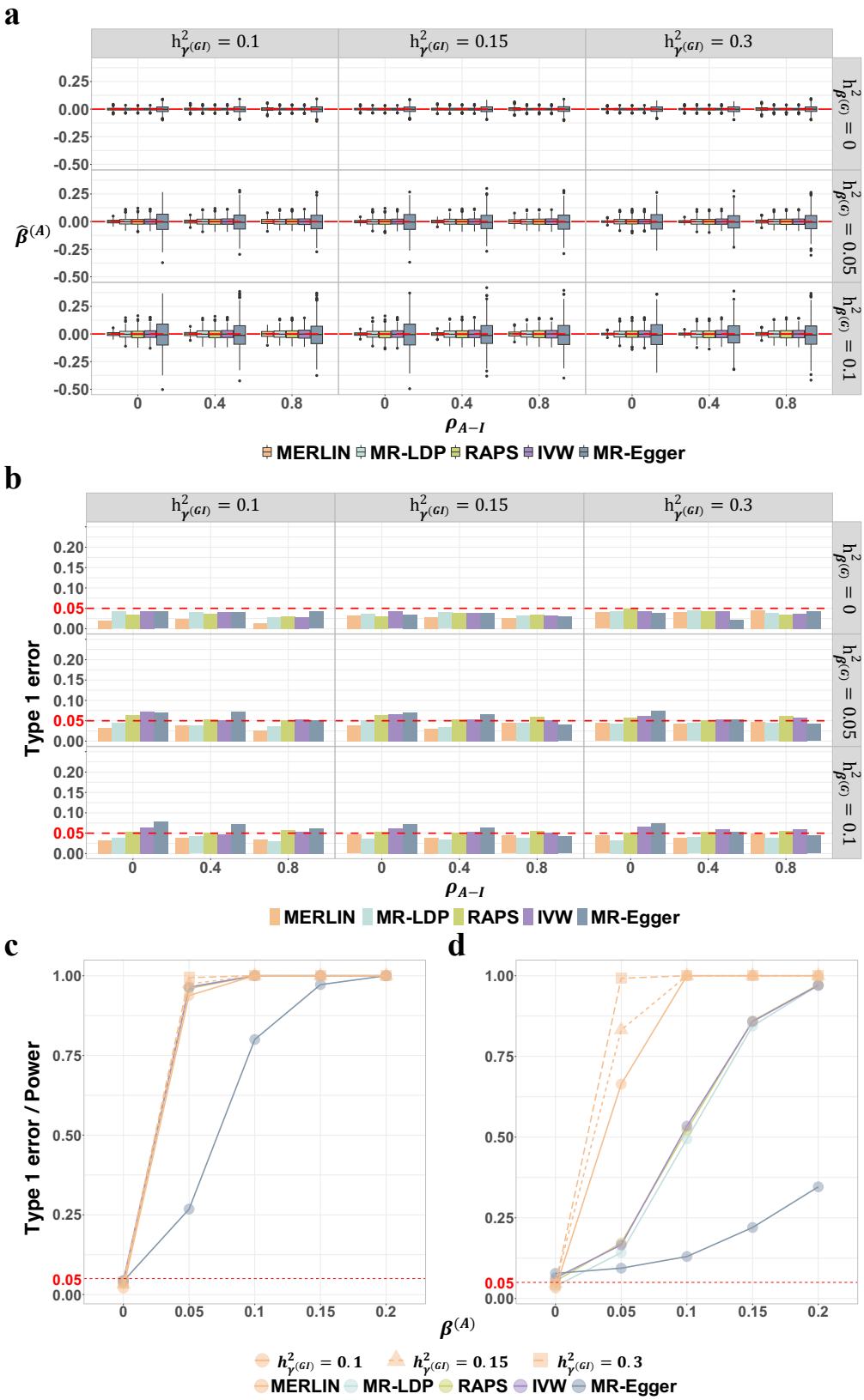

Figure S2. Simulation performance of MERLIN and standard MR methods for estimating the average causal effect ($\beta^{(A)}$). All simulations assumed a true interaction effect $\beta^{(I)} = 0$ and were based on 500 replicates. Standard MR methods included MR-LDP, RAPS, IVW, and MR-Egger. **(a)** Boxplots of average effect estimates ($\beta^{(A)}$) from MERLIN and standard MR methods. Performance is shown across varying proportions of exposure variance explained by

G×E effects $h^2_{\boldsymbol{\gamma}^{(GI)}} = 0.1, 0.15, 0.3$) and varying correlations between GWAS and GWIS instrument effects ($\rho_{(A-I)}$; x-axis within each plot: $0, 0.4, 0.8$). Scenarios are presented for true $\beta^{(A)} = 0$ with horizontal pleiotropy ($h^2_{\boldsymbol{\beta}^{(G)}} = 0, 0.05, 0.1$). Dashed red lines indicate the true $\beta^{(A)}$ values. **(b)** Type I error rates for testing $\beta^{(A)} = 0$, corresponding to the simulation conditions in panel (a). The dashed red line indicates the nominal 0.05 significance level. **(c, d)** Statistical power to detect $\beta^{(A)} \neq 0$ as a function of the true $\beta^{(A)}$ magnitude, assuming $\rho_{A-I} = 0$, with $h^2_{\boldsymbol{\beta}^{(G)}} = 0$ and 0.1, respectively. The plot also shows type I error at $\beta^{(A)} = 0$. Solid lines represent performance for all methods with $h^2_{\boldsymbol{\gamma}^{(GI)}} = 0.1$. Dashed lines illustrate MERLIN's performance with stronger G×E signals for exposure $h^2_{\boldsymbol{\gamma}^{(GI)}} = 0.15, 0.3$.

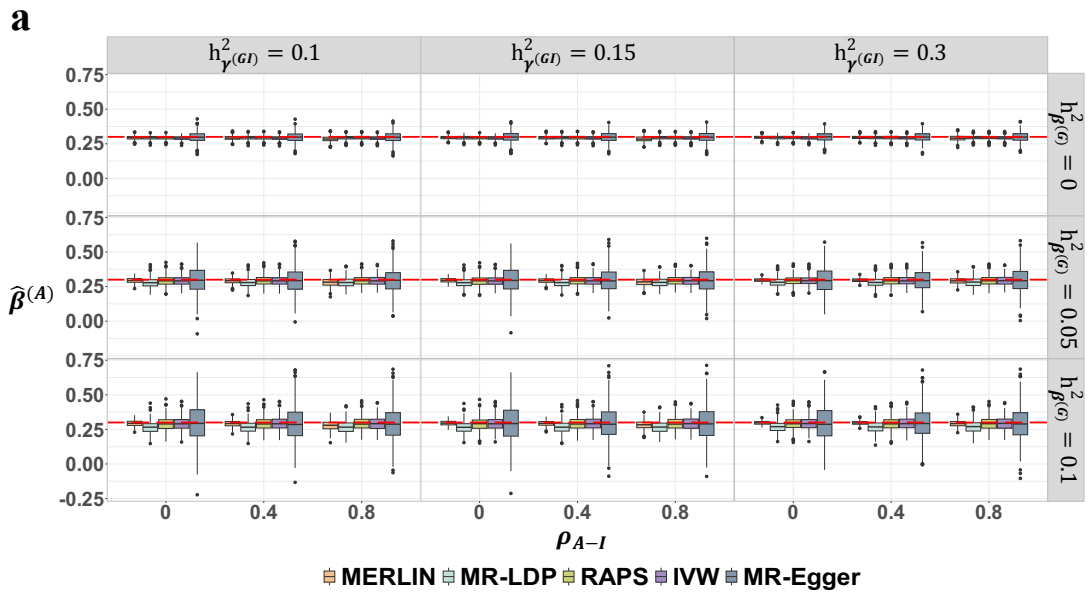

Figure S3. Simulation performance of MERLIN and standard MR methods for estimating the average causal effect ($\beta^{(A)}$). Simulation assumed a true average and interaction effect $\beta^{(A)} = 0.3$, $\beta^{(I)} = 0$ and were based on 500 replicates. Standard MR methods included MR-LDP, RAPS, IVW, and MR-Egger. Boxplots of average effect estimates ($\beta^{(A)}$) are shown across varying proportions of exposure variance explained by G×E effects ($h^2_{\gamma^{(GI)}} = 0.1, 0.15, 0.3$) and varying correlations between GWAS and GWIS instrument effects (ρ_{A-I}; x-axis within each plot: $0, 0.4, 0.8$) with horizontal pleiotropy ($h^2_{\beta^{(G)}} = 0, 0.05, 0.1$). Dashed red lines indicate the true $\beta^{(A)}$ values.

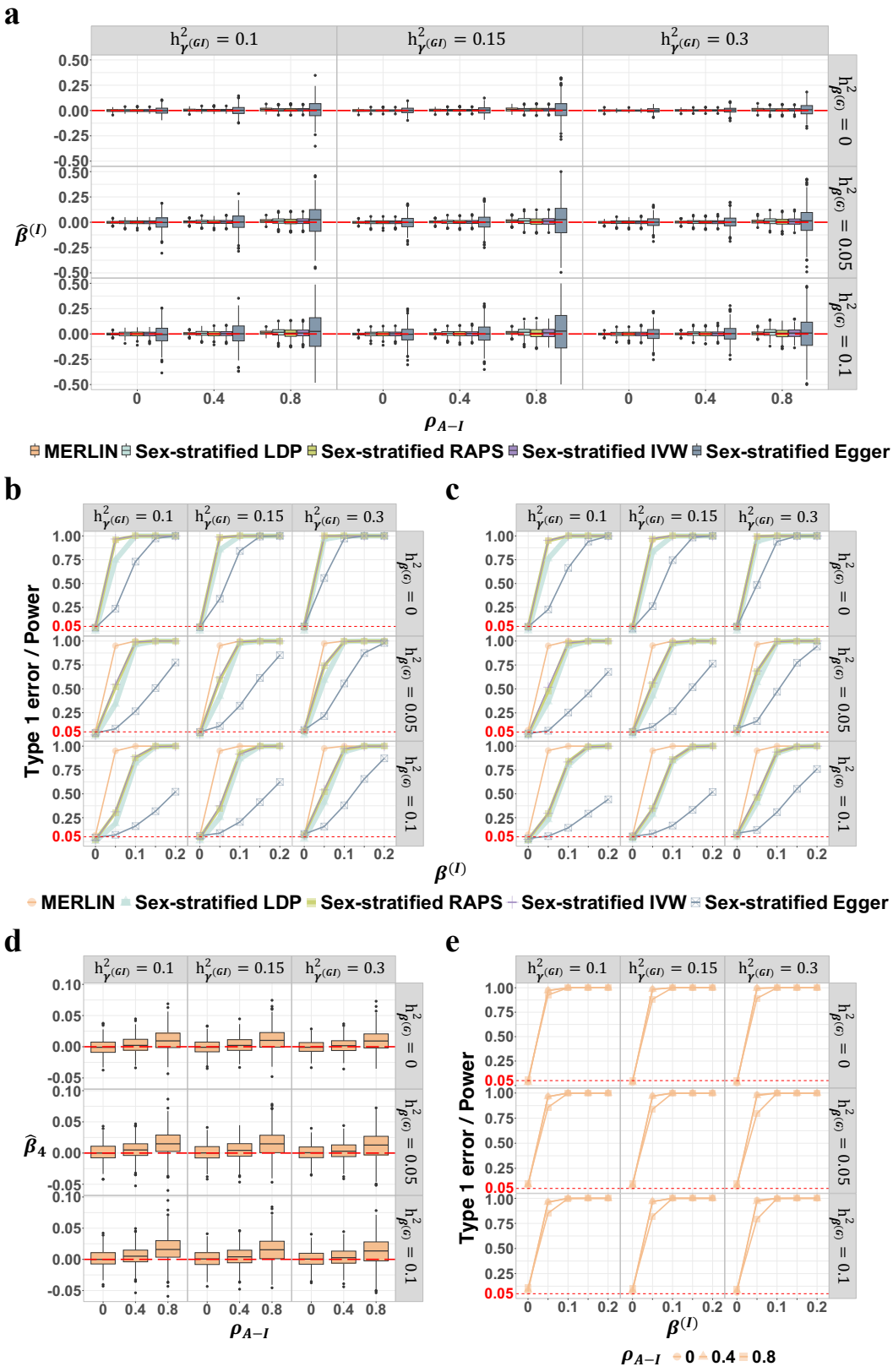

Figure S4. MERLIN accurately estimates interaction effects ($\beta^{(I)}$) and demonstrates superior statistical power in simulations. All simulations assumed a true average effect $\beta^{(A)} = 0.3$ and were based on 500 replicates. **(a)** Performance is shown across varying G×E signal strengths for exposure ($h^2_{\gamma(GI)}$; columns), correlations between GWAS and GWIS

instrument effects (ρ_{A-I}; x-axis of boxplots), and levels of horizontal pleiotropy ($h^2_{\boldsymbol{\beta}^{(G)}}$; rows). Dashed red lines indicate the true $\beta^{(I)} = 0$. Box plots show the unbiasedness of interaction effect estimates ($\beta^{(I)}$) from MERLIN and sex-stratified standard MR methods when the true $\beta^{(I)} = 0$. **(b)** Power comparison among MERLIN and sex-stratified standard MR methods in the detection of $\beta^{(I)} \neq 0$ for a discrete modifier in scenarios with no GWAS–GWIS correlation ($\rho_{A-I} = 0$) for varying horizontal pleiotropy ($h^2_{\boldsymbol{\beta}^{(G)}}$) and G×E signal strengths ($h^2_{\boldsymbol{\gamma}^{(GI)}}$). All methods control type I error (at true $\beta^{(I)} = 0$). **(c)** Power comparison as in (b) but in the presence of GWAS–GWIS correlation ($\rho_{A-I} = 0.4$). MERLIN shows a power advantage. **(d)** Box plots show the unbiasedness of $\beta^{(I)}$ (true $\beta^{(I)} = 0$) for a continuous modifier across conditions analogous to those in panel (a). **(e)** Statistical power of MERLIN in the detection of $\beta^{(I)} \neq 0$ for a continuous modifier. Power increases with true $\beta^{(I)}$ magnitude and G×E signal strength ($h^2_{\boldsymbol{\gamma}^{(GI)}}$), with type I error controlled, across various ρ_{A-I} values (shown for $h^2_{\boldsymbol{\beta}^{(G)}} = 0$).

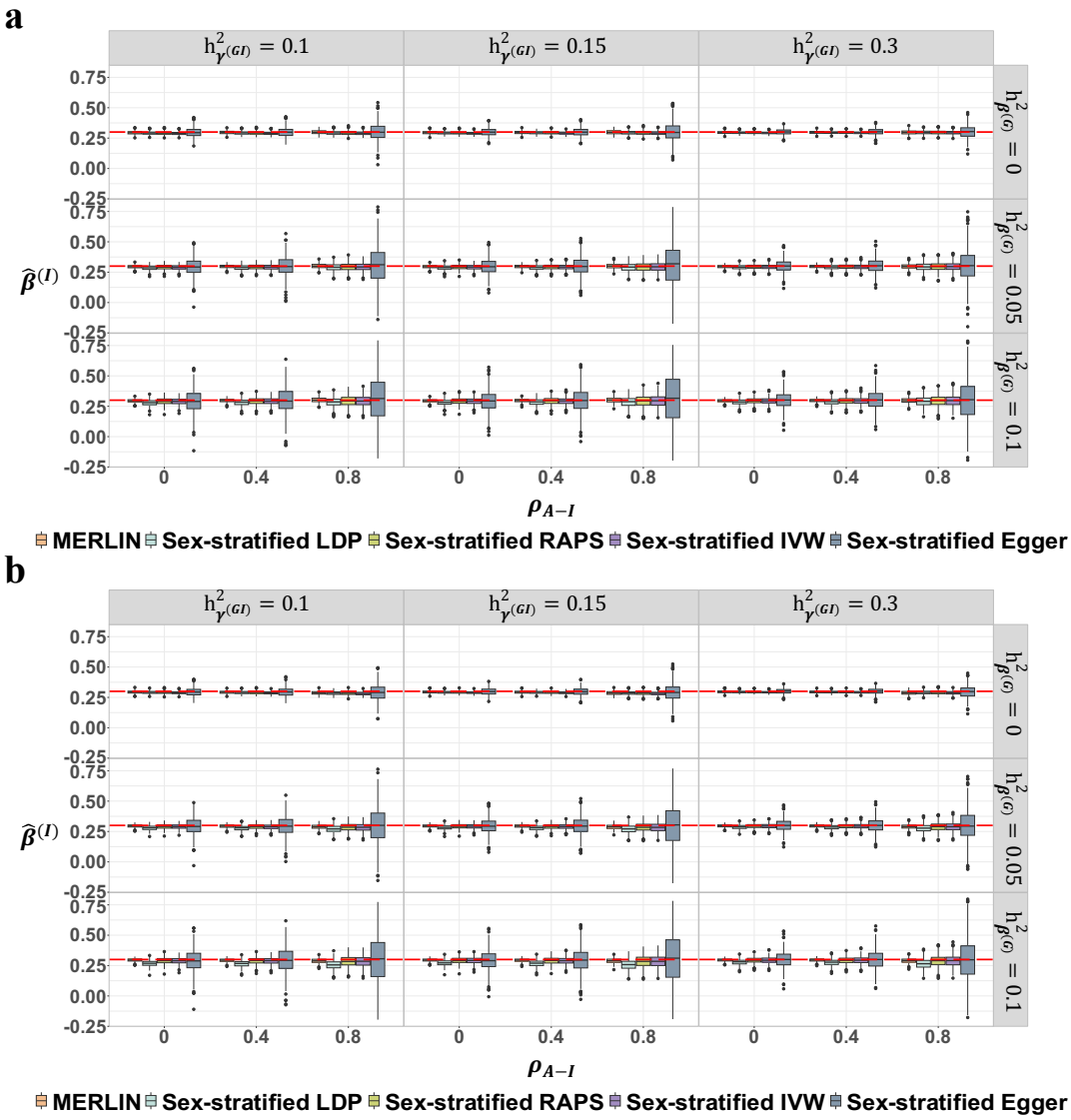

Figure S5. MERLIN accurately estimates interaction effects ($\beta^{(I)}$) in simulations. All simulations assumed a true interaction effect $\beta^{(I)} = 0.3$ and were based on 500 replicates. **(a)** Box plots show unbiasedness of interaction effect estimates ($\beta^{(I)}$) from MERLIN and sex-stratified standard MR methods when the true $\beta^{(A)} = 0.3$ for a discrete modifier. **(b)** Box plots show unbiasedness of interaction effect estimates ($\beta^{(I)}$) from MERLIN and sex-stratified standard MR methods when the true $\beta^{(A)} = 0$ for a discrete modifier. Dashed red lines in both (a) and (b) indicate the true $\beta^{(I)} = 0.3$. In both (a) and (b), performance is shown across varying G×E signal strengths for exposure ($h^2_{\gamma^{(GI)}}$; columns), correlations between GWAS and GWIS instrument effects (ρ_{A-I}; x-axis of boxplots), and levels of horizontal pleiotropy ($h^2_{\beta^{(G)}}$; rows).

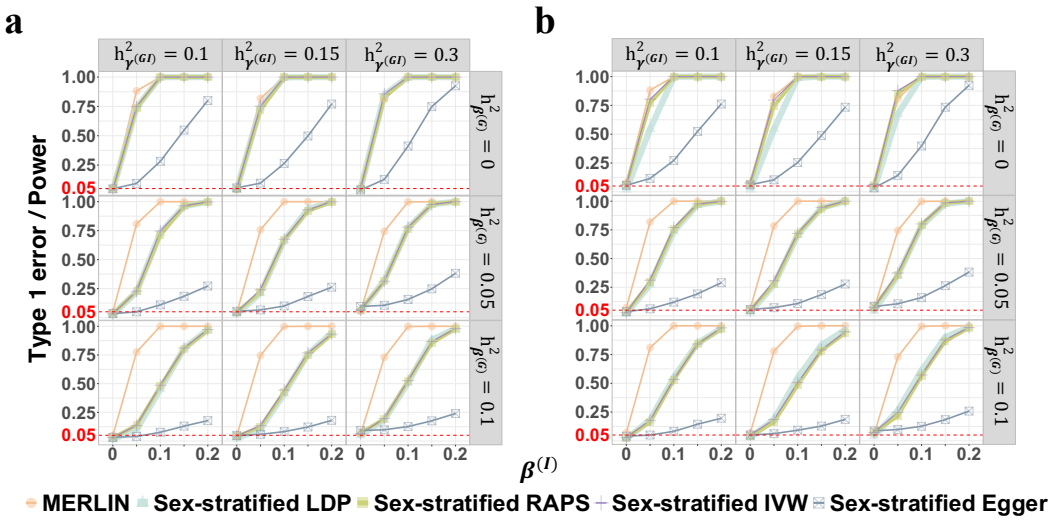

Figure S6. MERLIN demonstrates superior statistical power in simulations. (a) Simulations assumed a true average effect $\beta^{(A)} = 0$ and were based on 500 replicates. Power comparison between MERLIN and sex-stratified standard MR methods in detecting $\beta^{(I)} \neq 0$ for a discrete modifier in scenarios with no GWAS–GWIS correlation ($\rho_{A-I} = 0.8$) for varying horizontal pleiotropy ($h^2_{\beta^{(G)}}$) and G×E signal strengths ($h^2_{\gamma^{(GI)}}$). Both methods control type I error (at true $\beta^{(I)} = 0$). **(b)** Power comparison as in (a) but with $\beta^{(A)} = 0.3$ and in the presence of GWAS–GWIS correlation ($\rho_{A-I} = 0.8$). MERLIN shows a power advantage.

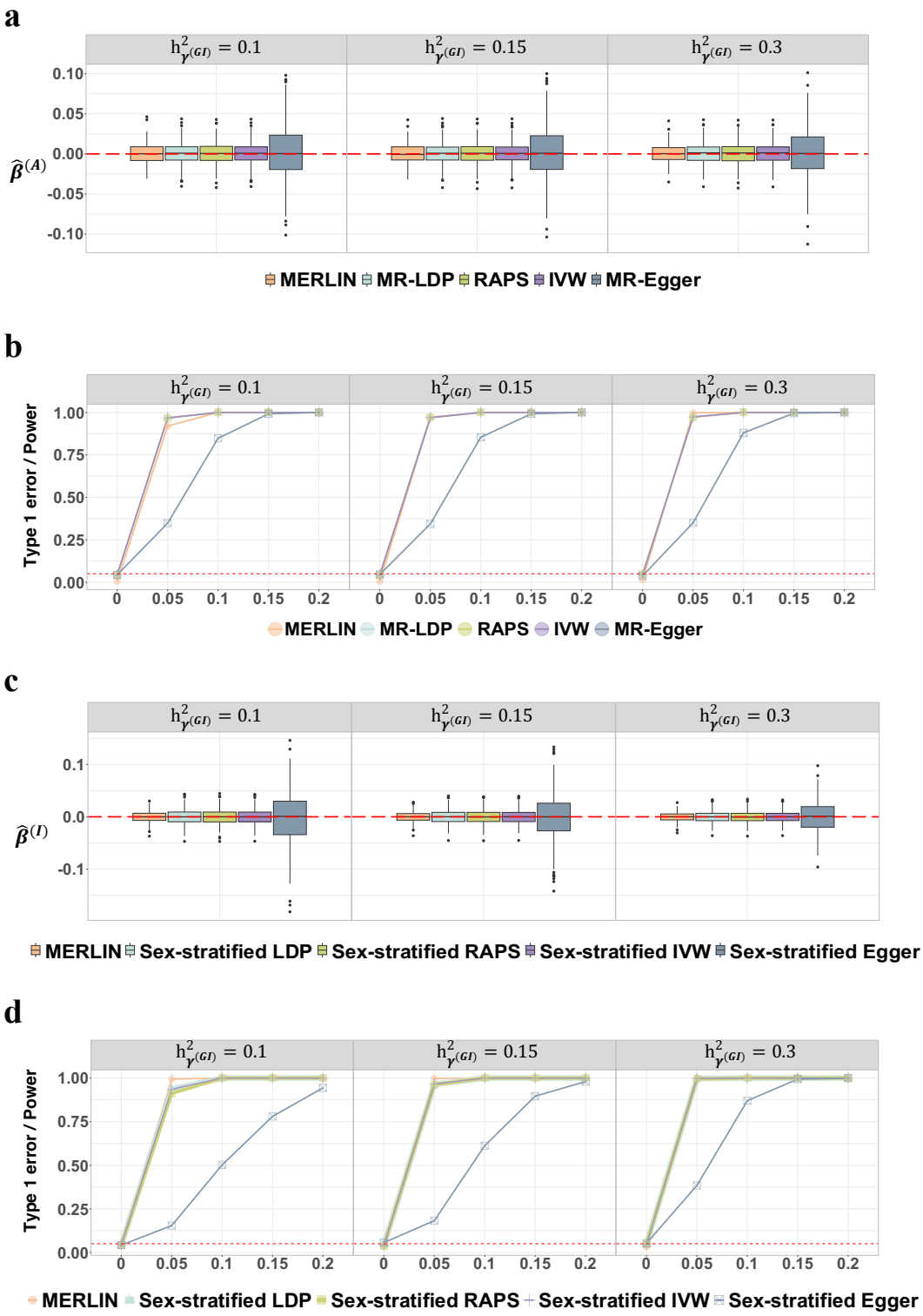

Figure S7. Simulation performance of MERLIN and standard MR methods for estimating the average causal effect ($\beta^{(A)}$) and the interaction effect ($\beta^{(I)}$) when the sample is imbalanced (the number of male samples is three times that of female samples). (a) Box plots show the average effect estimates ($\beta^{(A)}$) from MERLIN and standard MR methods when $\beta^{(I)} = 0$. Performance is shown across varying proportions of exposure variance explained by G × E effects ($h^2_{\gamma(GI)} = 0.1, 0.15, 0.3$) and fixed correlations between GWAS and GWIS instrument effects ($\rho_{A-I} = 0$). Scenarios are presented for true $\beta^{(A)} = 0$ with horizontal pleiotropy ($h^2_{\beta(G)} = 0$). Dashed red lines indicate the true $\beta^{(A)}$ values. **(b)** Statistical power to

detect $\beta^{(A)} \neq 0$ as a function of the true $\beta^{(A)}$ magnitude, while varying $h^2_{\boldsymbol{\gamma}^{(GI)}} = 0.1,\ 0.15,\ 0.3$, assuming $\rho_{A-I} = 0$, and with $h^2_{\boldsymbol{\beta}^{(G)}} = 0$. The plot also shows type I error at $\beta^{(A)} = 0$. **(c)** Boxplots of the interaction effect estimates ($\beta^{(I)}$) from MERLIN and sex-stratified standard MR methods when $\beta^{(A)} = 0$. The other settings are the same as in (a). **(d)** Power of the interaction effect estimates ($\beta^{(I)}$) from MERLIN and sex-stratified standard MR methods when $\beta^{(A)} = 0$. The other settings are the same as in (b)

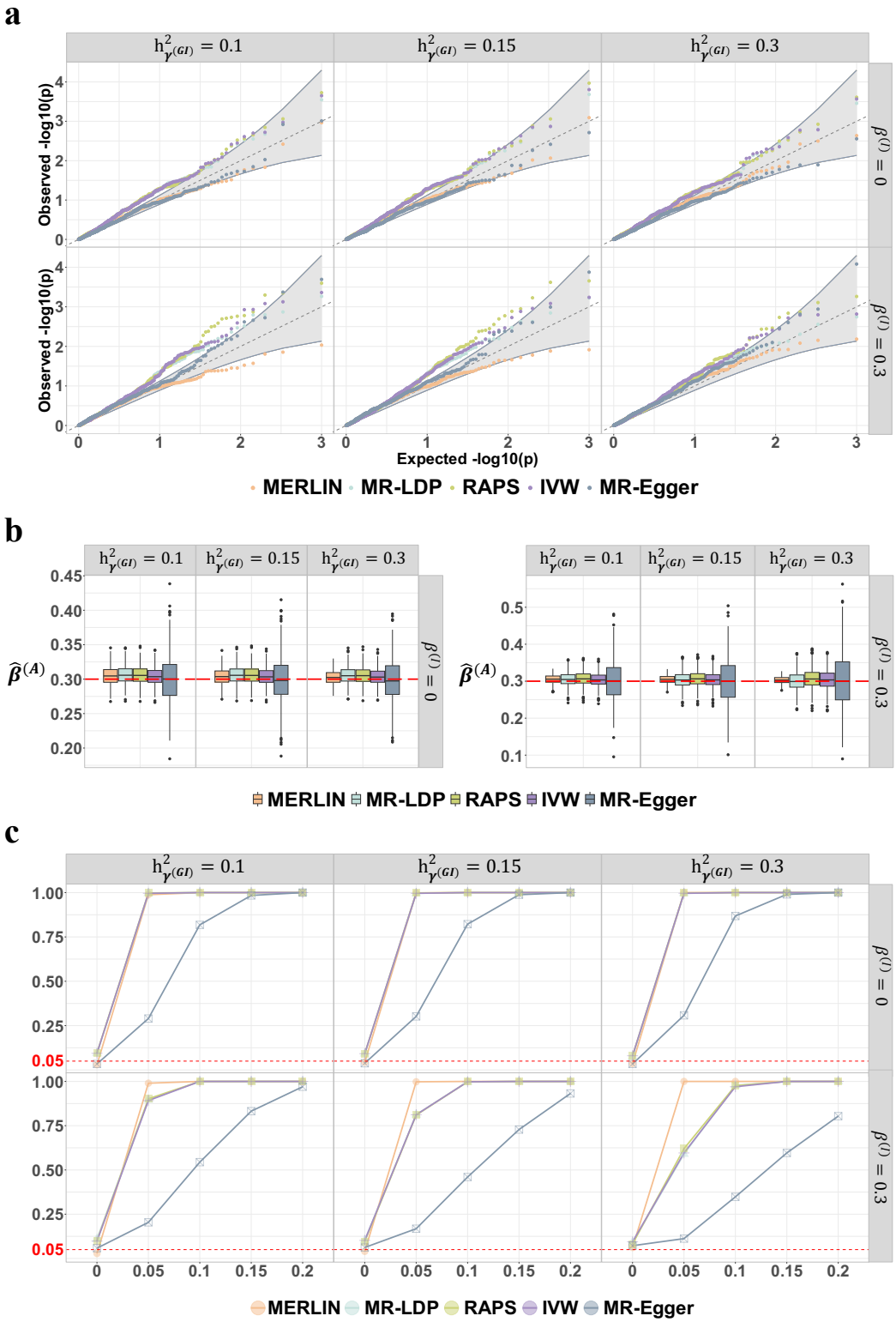

Figure S8. Simulation performance of MERLIN and standard MR methods for estimating the average causal effect ($\beta^{(A)}$) in sample overlaps. In all simulations, sample overlaps for males and females are each 20,000. **(a)** The QQ-plot of $\beta^{(A)}$ obtained by MERLIN and standard MR methods when $\beta^{(A)} = 0$ in scenarios with no GWAS–GWIS correlation ($\rho_{A-I} = 0$) for varying interaction effect ($\beta^{(I)}$) and G×E signal strengths ($h^2_{\gamma(GI)}$) **(b)** Estimates of $\beta^{(A)}$ comparison as in (a) but the true $\beta^{(A)} = 0.3$. MERLIN shows precise estimates. **(c)** The power of $\beta^{(A)}$ comparison as in (a), obtained by MERLIN and standard MR methods.

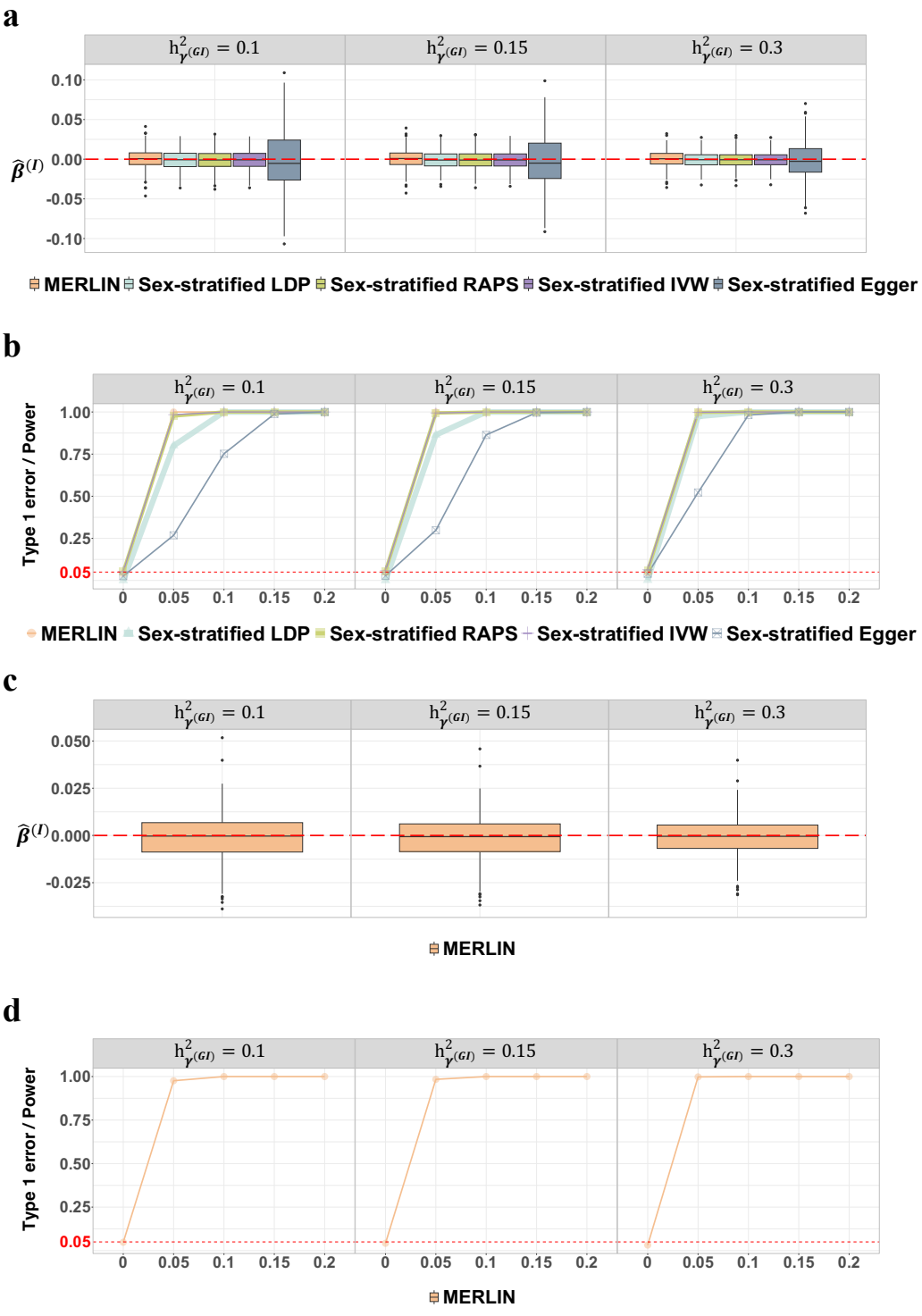

Figure S9. MERLIN accurately estimates interaction effects ($\beta^{(I)}$) and demonstrates superior statistical power in simulations when sample-overlaps for males and females are each 20,000. All simulations assumed a true average effect $\beta^{(A)} = 0.3$ and were based on 500 replicates. **(a)** Boxplots showing unbiasedness of interaction effect estimates ($\beta^{(I)}$) from MERLIN and sex-stratified standard MR methods when the true $\beta^{(I)} = 0$ for a discrete modifier. Performance is shown across varying G×E signal strengths for exposure ($h^2_{\gamma(GI)}$; columns) with no GWAS–GWIS correlation ($\rho_{A-I} = 0$) and fixed levels of horizontal pleiotropy ($h^2_{\beta(G)} = 0$; rows). Dashed red lines indicate the true $\beta^{(I)} = 0$. **(b)** Power comparison between MERLIN and sex-

stratified standard MR methods in detecting $\beta^{(I)} \neq 0$ for a discrete modifier in scenarios with no GWAS–GWIS correlation ($\rho_{A-I} = 0$) for fixed horizontal pleiotropy ($h^2_{\boldsymbol{\beta}^{(G)}} = 0$) and G×E signal strengths ($h^2_{\boldsymbol{\gamma}^{(GI)}}$). All methods control type I error (at true $\beta^{(I)} = 0$). **(c)** Boxplots showing that MERLIN provides unbiased $\beta^{(I)}$ estimates (true $\beta^{(I)} = 0$) for a continuous modifier across conditions analogous to those in panel (a). **(d)** Statistical power of MERLIN to detect $\beta^{(I)} \neq 0$ for a continuous modifier. Power increases with true $\beta^{(I)}$ magnitude and G×E signal strength ($h^2_{\boldsymbol{\gamma}^{(GI)}}$), with type I error controlled, across fixed $\rho_{A-I} = 0$ values (shown for $h^2_{\boldsymbol{\beta}^{(G)}} = 0$).

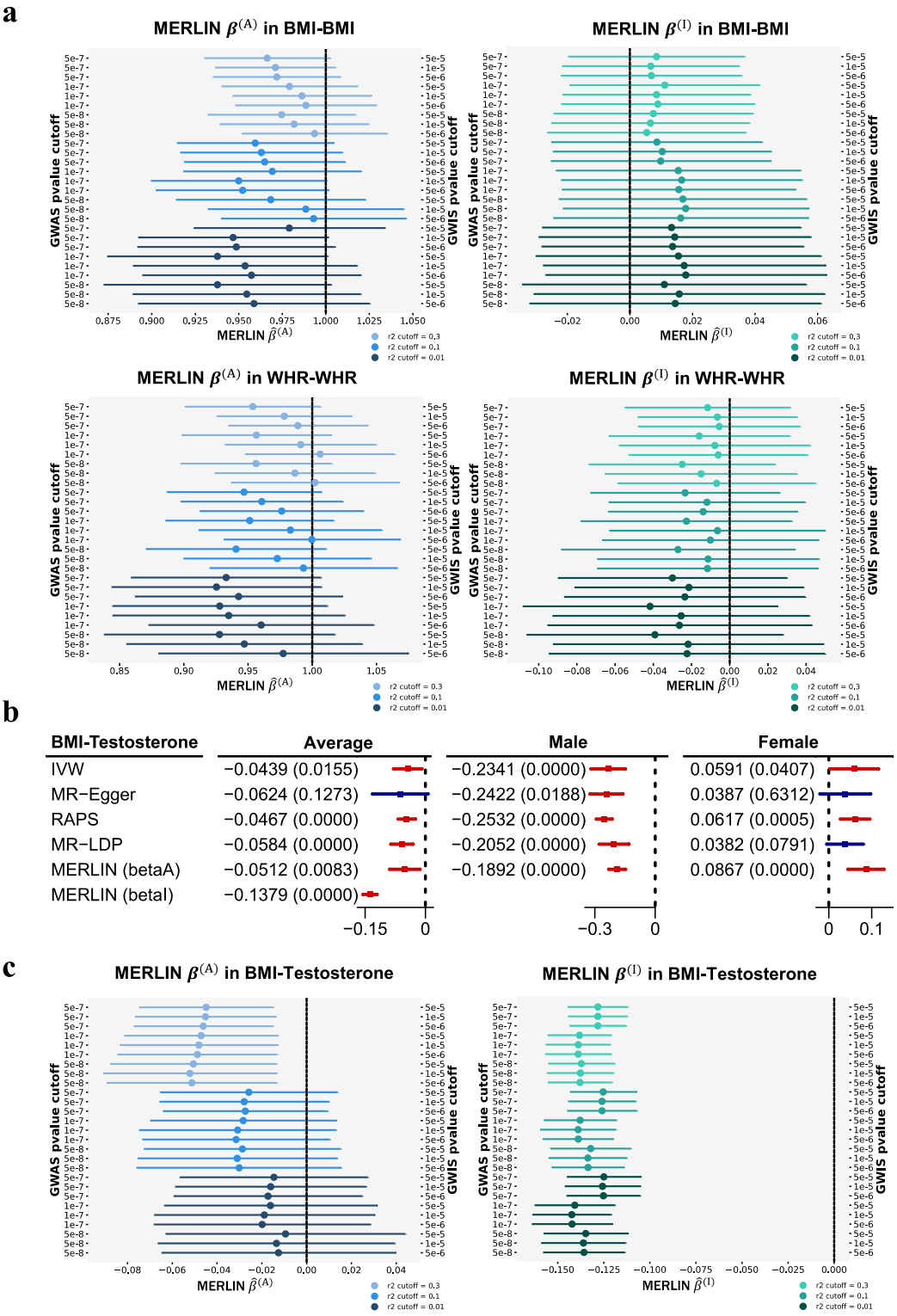

Figure S10. Sensitivity and sex-stratified analyses of MERLIN evaluated under negative and positive control scenarios. (a) Sensitivity analyses of MERLIN estimates across different IV selection thresholds for negative control. Across varying GWAS *P* value thresholds (5×10^{-8}, 1×10^{-7}, 5×10^{-7}), GWIS *P* value thresholds (5×10^{-6}, 1×10^{-5}, 5×10^{-5}), and LD clumping r^2 thresholds $(0.01, 0.1, 0.3)$, the estimates of $\beta^{(A)}$ consistently covered 1 and those of $\beta^{(I)}$ consistently covered 0, as expected under the negative control setting. **(b)** Positive

control analysis evaluating MERLIN's ability to detect known heterogeneity in the causal effect of body mass index (BMI) on testosterone level. MERLIN treats sex as a modifier, directly estimating the average effect ($\hat{\beta}^{(A)} = -0.0512, P = 0.0083$) and the sex-heterogeneity effect ($\hat{\beta}^{(I)} = -0.1379,\ P = 2.62 \times 10^{-54}$). Comparator methods provide overall average effects (column "Average") and sex-stratified estimates (columns "Male", "Female"). The MERLIN-derived sex-specific effects are consistent with these stratified results and published reports, demonstrating successful detection of the significant moderation by sex. All data are presented as effect estimates (*P* value) with 95% CIs. **(c)** Sensitivity analyses of MERLIN estimates across different IV selection thresholds for the positive control. Across varying GWAS *P* value thresholds ($5 \times 10^{-8},\ 1 \times 10^{-7},\ 5 \times 10^{-7}$), GWIS *P* value thresholds ($5 \times 10^{-6},\ 1 \times 10^{-5},\ 5 \times 10^{-5}$), and LD clumping r^2 thresholds $(0.01, 0.1, 0.3)$, the estimates of $\beta^{(I)}$ remained consistent and systematically deviated from the null, demonstrating significant sex-related heterogeneity. The estimates of $\beta^{(A)}$ were also relatively stable within the same r² threshold and became increasingly significant as the selection cutoffs grew more stringent (i.e., as the number of IVs decreased).

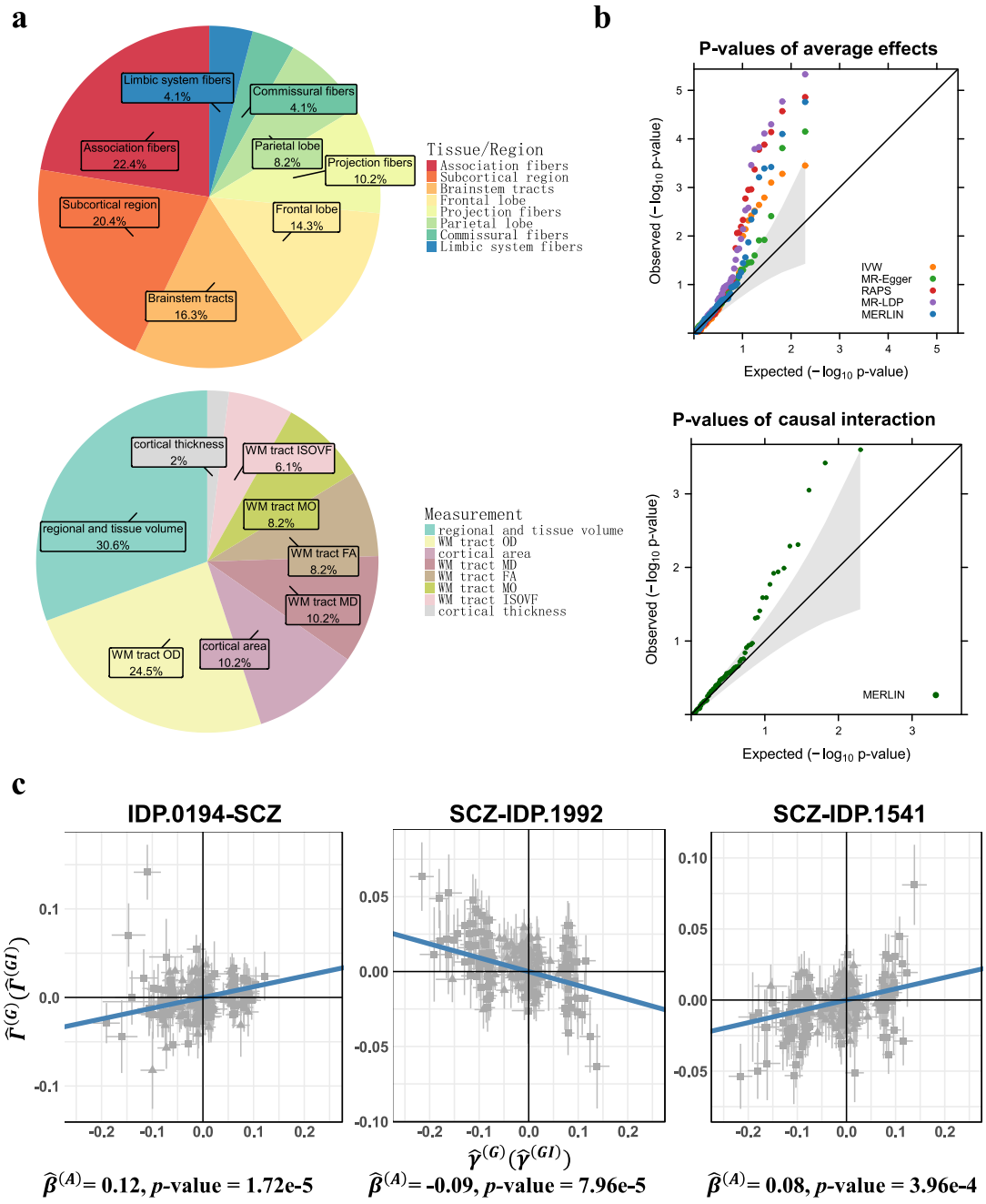

Figure S11. Genetic correlation, bidirectional MR testing, and MERLIN-based effect estimation between schizophrenia (SCZ) and brain imaging–derived phenotypes (IDPs). **(a)** Tissue and measurement distribution of brain IDPs after genetic correlation analyses. **(b)** Quantile–quantile plot of *P* values from the bidirectional Mendelian randomization analysis between SCZ and 49 brain IDPs. Average effects were estimated using IVW, MR-Egger, RAPS, MR-LDP, and MERLIN, whereas the heterogeneity effect could be estimated solely via MERLIN. IVW and MR-Egger were underpowered, whereas MERLIN provided more reliable estimates for both the average and heterogeneity effects. **(c)** MERLIN estimated significant average effects in IDP.0194-SCZ, SCZ-IDP.1992, and SCZ-IDP.1541. In the scatter plots, each data point corresponds to an individual SNP instrument. Each plot displays outcome GWAS ($\widehat{\boldsymbol{\Gamma}}^{(G)}$) and GWIS ($\widehat{\boldsymbol{\Gamma}}^{(GI)}$) effects versus exposure GWAS ($\widehat{\boldsymbol{\gamma}}^{(G)}$) and GWIS ($\widehat{\boldsymbol{\gamma}}^{(GI)}$) effects; the slope of the regression line corresponds to the average effect estimate, $\hat{\beta}^{(A)}$.

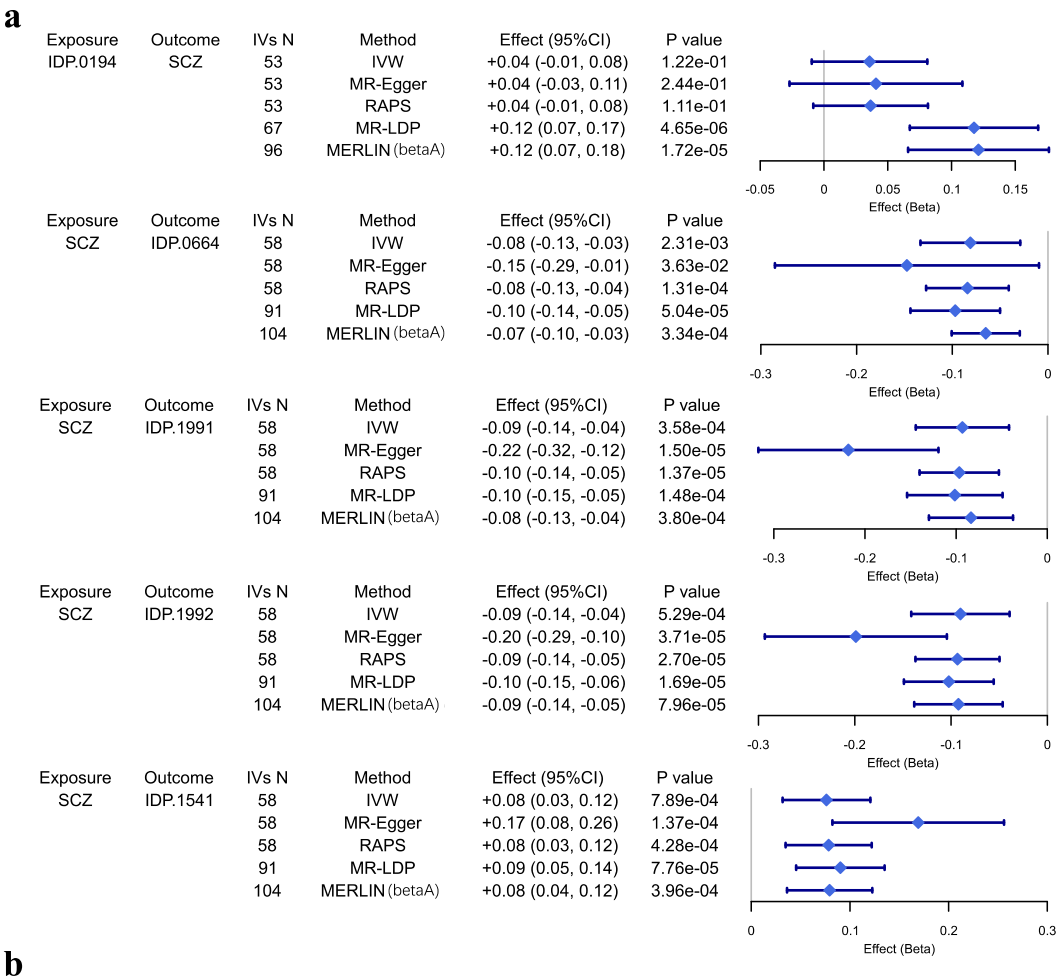

Figure S12. Summary of MERLIN-identified causal effects and sex-specific heterogeneity between schizophrenia (SCZ) and brain imaging–derived phenotypes (IDPs). (a) Forrest plots displaying five Bonferroni-corrected significant causal average effects identified by MERLIN involving schizophrenia (SCZ) and specific brain imaging-derived phenotypes (IDPs: IDP.0194, IDP.0664, IDP.1991, IDP.1992, IDP.1541). For each association, MERLIN's average effect estimates ($\hat{\beta}^{(A)}$) are shown alongside those from comparator MR methods (IVW, MR-Egger, RAPS, MR-LDP) and demonstrate directional consistency and comparable magnitudes. Effect estimates and 95% confidence intervals (CIs) are plotted; *P* values are provided. **(b)** Detailed analysis of two significant Bonferroni-corrected sex-heterogeneity effects identified by

MERLIN for the influence of SCZ on IDP.0664 and IDP.0015. Results are presented as effect estimate (*P* value) with 95% CIs. The “Average” column displays the average effect estimates from MERLIN and the comparator methods. The “Male” and “Female” columns display sex-specific causal effect estimates. For MERLIN, these are derived from its unified model ($\hat{\beta}_M = \hat{\beta}^{(A)} + \hat{\beta}^{(I)}$; $\hat{\beta}_F = \hat{\beta}^{(A)} - \hat{\beta}^{(I)}$, assuming male coded as +1 relative to the average effect); for IVW, MR-Egger, RAPS, and MR-LDP, these are results from traditional analyses stratified by sex.

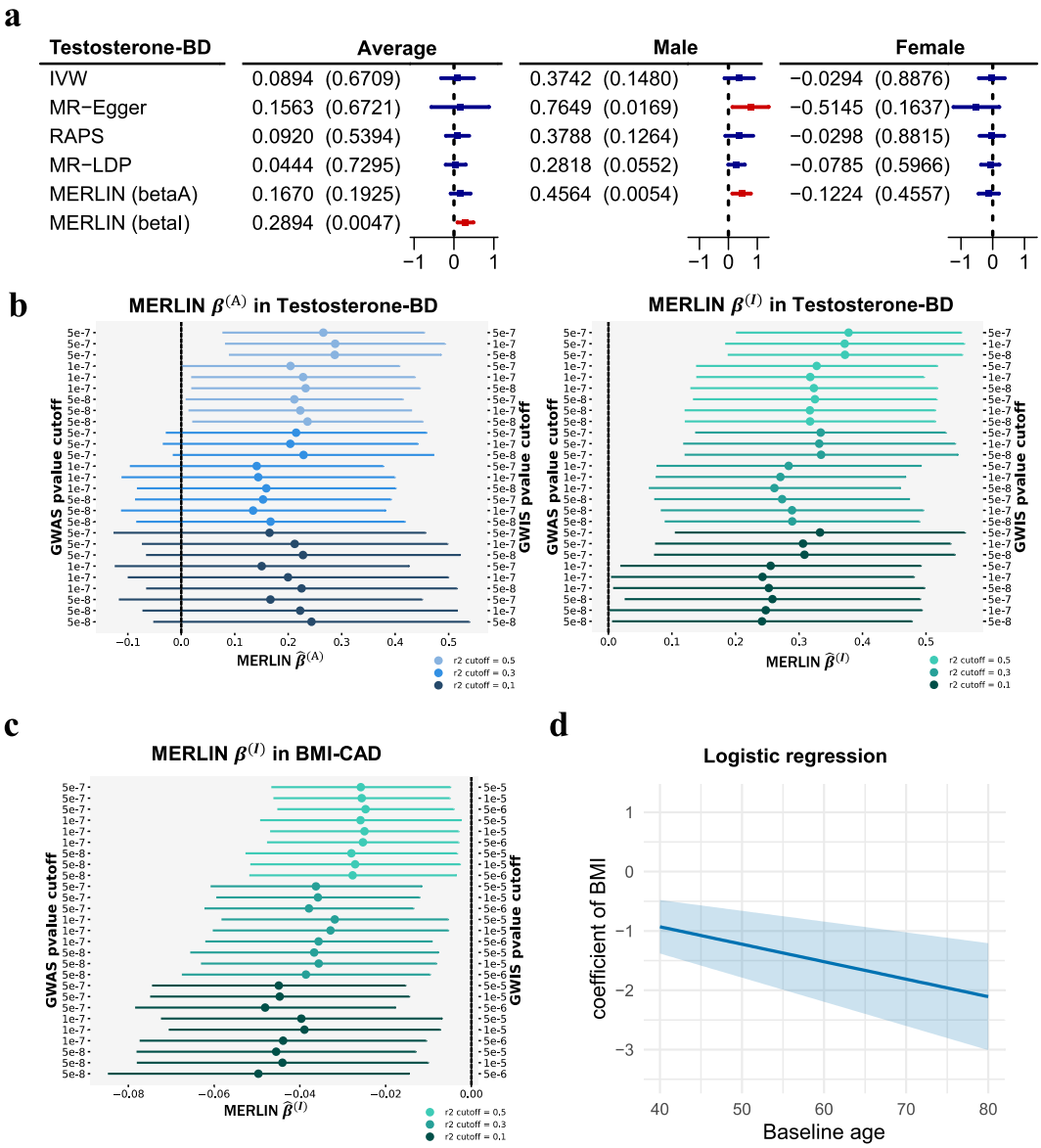

Figure S13. Sensitivity and age/sex-heterogeneity analyses for Testosterone–BD and BMI–CAD associations using MERLIN and comparator MR methods. (a) Comparative analysis of testosterone's effect on BD, with results presented as estimates (*P* value) and 95% CIs. The "Average" column shows average effect estimates: for MERLIN, this is its average effect parameter ($\hat{\beta}^{(A)} = 0.1670$, $P = 0.1925$); for IVW, MR-Egger, RAPS, and MR-LDP, these are their respective overall average effect estimates. MERLIN also directly estimated the causal interaction effect for BD ($\hat{\beta}^{(I)} = 0.2894$, $P = 0.0047$), which was statistically significant after Bonferroni correction for six traits tested. The "Male" and "Female" columns display sex-specific causal effect estimates: for MERLIN, these are derived from its unified model ($\hat{\beta}_M = \hat{\beta}^{(A)} + \hat{\beta}^{(I)}$; $\hat{\beta}_F = \hat{\beta}^{(A)} - \hat{\beta}^{(I)}$, assuming male coded as +1 relative to the average effect); for IVW, MR-Egger, RAPS, and MR-LDP, these are results from traditional analyses stratified by sex. **(b)** Sensitivity analyses of MERLIN estimates across different IV selection thresholds for Testosterone-BD. Across varying GWAS *P* value thresholds (5×10^{-8}, 1×10^{-7}, 5×10^{-7}), GWIS *P* value thresholds (5×10^{-8}, 1×10^{-7}, 5×10^{-7}), and LD clumping r² thresholds $(0.1, 0.3, 0.5)$, the estimates of $\beta^{(I)}$ remained consistent and systematically deviated from the null in both the original and alternative settings, demonstrating significant sex-related heterogeneity. The estimates of $\beta^{(A)}$ were also relatively stable and became increasingly significant as the selection cutoffs grew more stringent (i.e., as the number of IVs decreased). **(c)** Sensitivity analyses of MERLIN heterogeneity effect estimates across different IV selection thresholds for BMI-CAD. Across varying GWAS *P* value thresholds (5×10^{-8}, 1×10^{-7}, 5×10^{-7}), GWIS *P*

value thresholds ($5 \times 10^{-6},\ 1 \times 10^{-5},\ 5 \times 10^{-5}$), and LD clumping r^2 thresholds $(0.1, 0.3, 0.5)$, the estimates of $\beta^{(I)}$ remained consistent and systematically deviated from the null, demonstrating pronounced age-related heterogeneity. **(d)** Coefficient of BMI on the risk of CAD across baseline age from the logistic regression model. The graph illustrates the estimated coefficient (log odds) of standardized BMI on CAD risk as a function of the standardized baseline age, derived from a logistic regression model in the UKB data. The line represents the estimated coefficient, and the shaded area shows the 95% CI. The negative slope indicates that the positive association between BMI and CAD risk significantly attenuates with advancing age (interaction $\hat{\beta} = -0.0294,\ P < 3 \times 10^{-7}$).

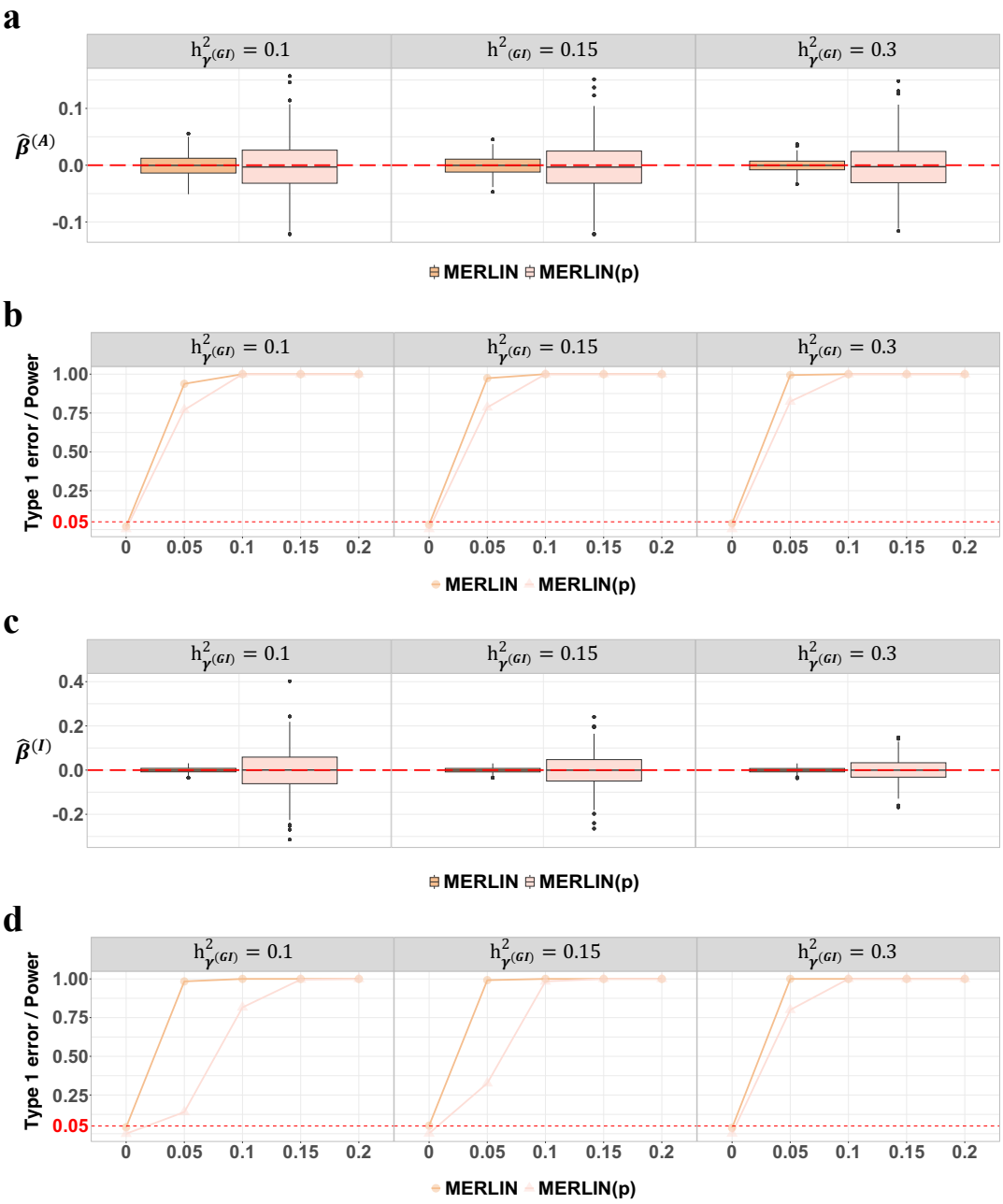

Figure S14. MERLIN accurately estimates the average effect ($\beta^{(A)}$) and interaction effects ($\beta^{(I)}$) and demonstrates superior statistical power in simulations. All simulations were based on 500 replicates. **(a)** Boxplots showing unbiasedness of the average effect estimates ($\beta^{(A)}$) from MERLIN and MERLIN(p) when the true $\beta^{(A)} = \beta^{(I)} = 0$. Performance is shown across varying G×E signal strengths for exposure ($h^2_{\gamma(GI)}$; columns) with no GWAS–GWIS correlation ($\rho_{A-I} = 0$) and fixed levels of horizontal pleiotropy ($h^2_{\beta(G)} = 0$; rows). Dashed red lines indicate the true $\beta^{(A)} = 0$. **(b)** Power comparison between MERLIN and MERLIN(p) detection of $\beta^{(A)} \neq 0$ (when $\beta^{(I)} = 0$) in scenarios with no GWAS–GWIS correlation ($\rho_{A-I} = 0$) for fixed horizontal pleiotropy ($h^2_{\beta(G)} = 0$) and G×E signal strengths ($h^2_{\gamma(GI)}$). Both methods control type I error (at true $\beta^{(I)} = 0$). **(c)** Box plots show MERLIN and MERLIN(p) estimates of $\beta^{(I)}$ for the same scenario as (a). **(d)** Statistical power of MERLIN in the detection of $\beta^{(I)} \neq 0$ for the same scenario as (b)

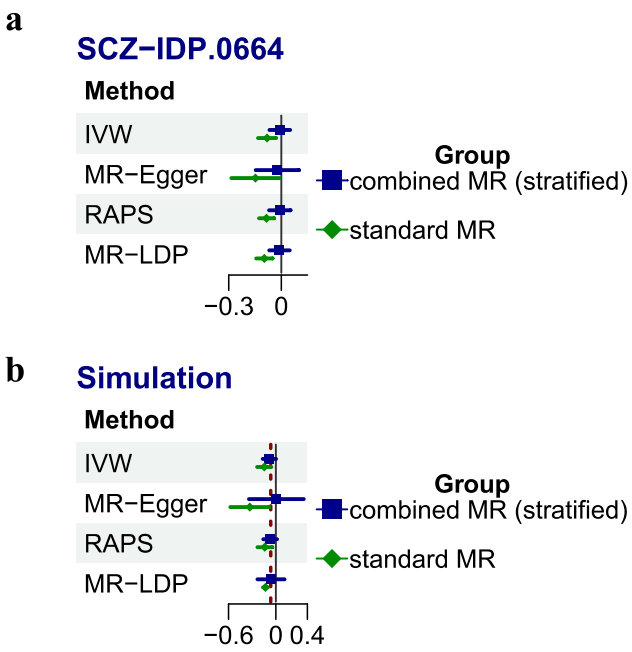

Figure S15. Bias in standard MR average effect estimation under heterogeneity for the SCZ → IDP.0664 association. (a) In the real-data analysis, the average effect ($\beta^{(A)}$) estimates from standard MR (green) diverged from the more robust sex-stratified combined MR estimates (blue). **(b)** A targeted simulation confirmed this bias. Boxplots of estimates from standard MR (green) show systematic bias away from the true value (dashed line, $\beta^{(A)} = -0.065$), while the sex-stratified combined MR estimates (blue) are closer to the true average effect.