

Summary statistic analyses do not correct confounding bias: Appendix

John B. Holmes^{*1}, Doug Speed^{2,3}, and David J. Balding^{1,2}

¹Melbourne Integrative Genomics, School of Mathematics and Statistics, University of Melbourne, Australia.

²Aarhus Institute of Advanced Studies (AIAS), Aarhus University, Denmark.

³UCL Genetics Institute, University College London, United Kingdom.

January 25, 2019

1 Association between LD score and confounding

To simplify the presentation we only consider the LD score as defined in LDSC, and not the weighted version used by SumHer, although similar issues do arise for SumHer. We assume without loss of generality that \mathbf{X} is orthonormal with $< n$ columns, and \mathbf{Z} is standardised. Let $\mathbf{Z} = \mathbf{XB} + \mathbf{E}$. Then

$$\text{Cor}(\mathbf{Z}) = \text{Cov}(\mathbf{Z}) = \text{Cov}(\mathbf{XB} + \mathbf{E}) = \mathbf{B}'\text{Cov}(\mathbf{X})\mathbf{B} + \text{Cov}(\mathbf{E}) = \mathbf{B}'\mathbf{B} + \text{Cov}(\mathbf{E}), \quad (1)$$

$$\text{Cov}(\mathbf{Z}_j, \mathbf{X}\boldsymbol{\beta}) = \text{Cov}(\mathbf{XB}_j, \mathbf{X}\boldsymbol{\beta}) = \mathbf{B}'_j\text{Cov}(\mathbf{X})\boldsymbol{\beta} = \mathbf{B}'_j\boldsymbol{\beta}. \quad (2)$$

Hence $(\mathbf{B}'_j\boldsymbol{\beta})^2 = a_j$ and we will write $\mathbf{B}'_j\boldsymbol{\beta} = \sqrt{a_j}$ which can be positive or negative. Now decompose \mathbf{B}_j into a multiple of $\boldsymbol{\beta}$ plus a vector $\hat{\boldsymbol{\epsilon}}_j$ that is orthogonal to $\boldsymbol{\beta}$. Then the coefficient of $\boldsymbol{\beta}$ is $\boldsymbol{\beta}'\mathbf{B}_j/\boldsymbol{\beta}'\boldsymbol{\beta} = \sqrt{a_j}/\boldsymbol{\beta}'\boldsymbol{\beta}$. Substituting (1) into the LD score expression, we obtain

$$\begin{aligned} \sum_i r_{ij}^2 &= \sum_i \text{Cor}(\mathbf{Z})_{ij}^2 = \sum_i (\mathbf{B}'_i\mathbf{B}_j + \text{Cov}(\mathbf{E})_{ij})^2 = \sum_i ((\boldsymbol{\beta}\sqrt{a_i}/\boldsymbol{\beta}'\boldsymbol{\beta} + \hat{\boldsymbol{\epsilon}}_i)'(\boldsymbol{\beta}\sqrt{a_j}/\boldsymbol{\beta}'\boldsymbol{\beta} + \hat{\boldsymbol{\epsilon}}_j) + \text{Cov}(\mathbf{E})_{ij})^2 \\ &= \sum_i (\sqrt{a_i a_j}/\boldsymbol{\beta}'\boldsymbol{\beta} + \hat{\boldsymbol{\epsilon}}_i'\hat{\boldsymbol{\epsilon}}_j + \text{Cov}(\mathbf{E})_{ij})^2 \\ &= \frac{a_j}{(\boldsymbol{\beta}'\boldsymbol{\beta})^2} \sum_i a_i + 2\frac{\sqrt{a_j}}{\boldsymbol{\beta}'\boldsymbol{\beta}} \sum_i \sqrt{a_i}(\hat{\boldsymbol{\epsilon}}_i'\hat{\boldsymbol{\epsilon}}_j + \text{Cov}(\mathbf{E})_{ij}) + \sum_i (\hat{\boldsymbol{\epsilon}}_i'\hat{\boldsymbol{\epsilon}}_j + \text{Cov}(\mathbf{E})_{ij})^2. \end{aligned} \quad (3)$$

The middle line above uses the orthogonality (by definition) of $\boldsymbol{\beta}$ and ϵ_j . As Eq. 1 in the main text does not accommodate dependence between a_j and $\sum_i r_{ij}^2$, resulting estimates of h_{SNP}^2 are in general

^{*}corresponding author, email: john.holmes@unimelb.edu.au

biased. As $\sqrt{a_i}$, and $\hat{\epsilon}_i' \hat{\epsilon}_j + \text{Cov}(\mathbf{E})_{ij}$ can be either positive or negative, the sum associated with $\sqrt{a_j}$ will tend to cancel so that LD score is almost always positively correlated with a_j , which in LDSC will inflate estimates of h_{SNP}^2 . As $\sum_i a_i = \sum_i \text{Cov}(\mathbf{Z}_i, \mathbf{X}\boldsymbol{\beta})^2$, the magnitude of inflation increases with the square correlations between the genotypes \mathbf{Z} and the ignored confounding effect $\mathbf{X}\boldsymbol{\beta}$ as well as with the proportion of phenotypic variation attributed to the confounding effect.

To aid intuition about this effect, consider the widespread and successful use of principal components of SNP genotypes to adjust for confounding due to population structure. In order to maximise variation explained, the first principal component usually has high loadings from sets of SNPs in high LD and it usually captures a large component of confounding, supporting our claim of an association between LD and confounding. If the stratification is due to ignored sub-populations, SNPs with greater allele frequency differences between sub-populations will have greater induced LD, so that higher LD score will be associated with the SNPs more predictive of sub-population membership.

To illustrate the impact of the LD/confounding relationship on estimation of h_{SNP}^2 , we extracted the confounding effects $\mathbf{X}\boldsymbol{\beta}$ from our simulations, and treated their standardised values as the phenotype in a GWAS analysis. The resulting estimate of h_{SNP}^2 was multiplied by 0.1 to have the same scale as the confounding component in a $C1$ phenotype. The positive relationship between LD and confounding generates positive estimates of h_{SNP}^2 (Figure S10), although these estimates tend to be lower than the observed inflation in h_{SNP}^2 estimates in the original analyses. This is consistent with the expected cross-product of the confounding effect test statistic and confounding-free phenotype test statistic also showing a positive relationship with LD score.

2 Approximate expectations of association statistics

We restate our general phenotype model

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\alpha} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_e^2 \mathbf{I}), \quad (4)$$

here making no distributional assumption for $\boldsymbol{\beta}$ and assuming that columns of \mathbf{X} are standardised.

2.1 Single-SNP analysis with no covariates fitted and \mathbf{Z} known

In single-SNP regression of a quantitative phenotype,

$$T_j^2 = \frac{(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})' \mathbf{Z}_j (\mathbf{Z}_j' \mathbf{Z}_j)^{-1} \mathbf{Z}_j' (\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})}{(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})' (\mathbf{I} - \mathbf{Z}_j (\mathbf{Z}_j' \mathbf{Z}_j)^{-1} \mathbf{Z}_j') (\mathbf{y} - \bar{\mathbf{y}}\mathbf{1}) / (n-2)}. \quad (5)$$

Following [1], we consider the numerator and denominator of (5) separately. Some notation: $\mathbf{J} = \mathbf{1}\mathbf{1}'$ is a matrix of ones, so that $\mathbf{J}^2 = n\mathbf{J}$, and Tr denotes the trace of a matrix. In particular if \mathbf{a} is a vector, $\text{Tr}(\mathbf{a}\mathbf{a}') = \mathbf{a}'\mathbf{a}$. We require

$$\begin{aligned}
\text{E}[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'] &= (\mathbf{I} - \mathbf{J}/n)\text{E}[\mathbf{y}\mathbf{y}'](\mathbf{I} - \mathbf{J}/n)' \\
&= (\mathbf{I} - \mathbf{J}/n)(\mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2\mathbf{I} + \mathbf{X}\beta\beta'\mathbf{X}')(\mathbf{I} - \mathbf{J}/n)' \\
&= \mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2(\mathbf{I} - 2\mathbf{J}/n + \mathbf{J}\mathbf{J}/n^2) + \mathbf{X}\beta\beta'\mathbf{X}' \\
&= \mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2(\mathbf{I} - \mathbf{J}/n) + \mathbf{X}\beta\beta'\mathbf{X}'
\end{aligned} \tag{6}$$

with the final two equalities following from the standardisation of \mathbf{Z} , which implies $\mathbf{Z}'\mathbf{J} = \mathbf{0}_{m \times n}$ and standardisation of \mathbf{X} implying $\mathbf{X}'\mathbf{J} = \mathbf{0}_{p \times n}$. Using (7), the expected value of the numerator of (5), which following ANOVA conventions we call SSR_j , is

$$\begin{aligned}
\text{E}[SSR_j|\mathbf{Z}, \mathbf{X}] &= \text{E}[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{Z}_j'(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})] \\
&= \text{Tr}(\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\text{E}[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})']) \\
&= \text{Tr}(\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\{\mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2(\mathbf{I}_n - \mathbf{J}/n) + \mathbf{X}\beta\beta'\mathbf{X}'\}) \\
&= \text{Tr}(\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\{\mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2\mathbf{I}_n + \mathbf{X}\beta\beta'\mathbf{X}'\}) \\
&= \text{Tr}((\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{Z}_j + \sigma_e^2(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}(\mathbf{Z}_j'\mathbf{Z}_j) + (\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{Z}_j) \\
&= (\mathbf{Z}_j'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{Z}_j + \mathbf{Z}_j'\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{Z}_j)/(n-1) + \sigma_e^2.
\end{aligned} \tag{8}$$

Note that $\mathbf{Z}_j'\mathbf{Z} = (n-1)\hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{Z})$, and if \mathbf{X} is standardised, then $\mathbf{Z}_j'\mathbf{X} = (n-1)\hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X})$. Since Σ is diagonal, we can write

$$\begin{aligned}
\text{E}[SSR_j|\mathbf{Z}, \mathbf{X}] &= (n-1)\left\{\sum_{i=1}^m \hat{r}_{ij}^2\sigma_i^2 + \left(\sum_{k=1}^p \hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k)\beta_k\right)^2\right\} + \sigma_e^2 \\
&= \sigma_e^2 + (n-1)\sigma_j^2 + (n-1)\sum_{i \neq j} \hat{r}_{ij}^2\sigma_i^2 + (n-1)\left(\sum_{k=1}^p \hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k)\beta_k\right)^2
\end{aligned} \tag{9}$$

where $\hat{r}_{ij} = \hat{\text{Cor}}(\mathbf{Z}_i, \mathbf{Z}_j)$. Write SSE_j for $n-2$ times the denominator of (5). Then

$$\begin{aligned}
\text{E}[SSE_j|\mathbf{Z}, \mathbf{X}] &= \text{E}[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'(\mathbf{I} - \mathbf{Z}_j(\mathbf{Z}_j'\mathbf{Z}_j)^{-1}\mathbf{Z}_j')(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})] \\
&= \text{E}[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})] - \text{E}[SSR_j|\mathbf{Z}, \mathbf{X}] \\
&= \text{Tr}(\text{E}[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})']) - \text{E}[SSR_j|\mathbf{Z}, \mathbf{X}] \\
&= \text{Tr}(\mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2(\mathbf{I} - \mathbf{J}/n) + \mathbf{X}\beta\beta'\mathbf{X}') - \text{E}[SSR_j|\mathbf{Z}, \mathbf{X}]
\end{aligned}$$

$$\begin{aligned}
&= \text{Tr}(\mathbf{Z}'\mathbf{Z}\Sigma) + (n-1)\sigma_e^2 + \text{Tr}(\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) - \text{E}[SSR_j|\mathbf{Z}, \mathbf{X}] \\
&= (n-1)\left\{ \sum_{i=1}^m \sigma_j^2 + \sigma_e^2 + \boldsymbol{\beta}'\hat{\text{Var}}[\mathbf{X}]\boldsymbol{\beta} \right\} - \sigma_e^2 - (n-1)\left\{ \sum_{i=1}^m \hat{r}_{ij}^2 \sigma_i^2 - \left(\sum_{k=1}^p \hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k) \boldsymbol{\beta}_k \right)^2 \right\} \\
&= (n-2)\sigma_e^2 + (n-1)\left\{ \sum_{i \neq j} (1 - \hat{r}_{ij}^2) \sigma_i^2 + \sum_{k=1}^p \sum_{l=1}^p (\hat{\text{Var}}[\mathbf{X}]_{kl} - \hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k) \hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_l)) \boldsymbol{\beta}_k \boldsymbol{\beta}_l \right\} \quad (10)
\end{aligned}$$

2.2 Single-SNP analysis with no covariates fitted and \mathbf{Z} unknown

The reported results from a GWAS typically do not include the sample SNP correlation matrix, which is required for the expectations in section 2.1. Normally, it is replaced by an estimate from a reference panel. Rather than use \hat{r}_{ij}^2 directly, [1] used a transformation to reduce bias: $r_{ij}^2 = ((n-1)\hat{r}_{ij}^2 - 1)/(n-2)$. Substituting into (9) for SNP-SNP and SNP-covariate correlations, we obtain

$$\begin{aligned}
\text{E}[SSR_j|\mathbf{X}] &= \sigma_e^2 + (n-1)\sigma_j^2 + (n-1) \sum_{i \neq j} \left\{ \frac{(n-2)r_{ij}^2}{n-1} + \frac{1}{n-1} \right\} \sigma_j^2 + (n-1) \left(\sum_{k=1}^p \hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k) \boldsymbol{\beta}_k \right)^2 \\
&= \sigma_e^2 + \sum_{i=1}^m \sigma_j^2 + (n-2)\sigma_j^2 + (n-2) \sum_{i \neq j} r_{ij}^2 \sigma_i^2 + (n-1)\sigma_e^2 \hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}\boldsymbol{\beta})^2 \\
&= \sigma_e^2 + \sum_{i=1}^m \sigma_j^2 + (n-2)\sigma_j^2 + (n-2) \sum_{i \neq j} r_{ij}^2 \sigma_i^2 + (n-1)\sigma_e^2 \left\{ \frac{(n-2)\text{Cor}(\mathbf{Z}_j, \mathbf{X}\boldsymbol{\beta})^2}{n-1} + \frac{1}{n-1} \right\} \\
&= \sigma_y^2 + \sigma_c^2 + (n-2)\sigma_j^2 + (n-2) \sum_{i \neq j} r_{ij}^2 \sigma_i^2 + (n-2) \left(\sum_{k=1}^p \text{Cor}(\mathbf{Z}_j, \mathbf{X}_k) \boldsymbol{\beta}_k \right)^2. \quad (11)
\end{aligned}$$

For large n ,

$$\text{E}[SSR_j|\mathbf{X}] \approx \sigma_y^2 + \sigma_c^2 + n\sigma_j^2 + n \sum_{i \neq j} r_{ij}^2 \sigma_i^2 + n \sum_{k=1}^p \sum_{l=1}^p \text{Cor}(\mathbf{Z}_j, \mathbf{X}_k) \text{Cor}(\mathbf{Z}_j, \mathbf{X}_l) \boldsymbol{\beta}_k \boldsymbol{\beta}_l. \quad (12)$$

$$\begin{aligned}
\text{E}[SSE_j|\mathbf{X}] &= \text{E}[\text{E}[SSE_j|\mathbf{Z}, \mathbf{X}]] = \text{E}[\text{E}[SST|\mathbf{Z}, \mathbf{X}]] - \text{E}[\text{E}[SSR_j|\mathbf{Z}, \mathbf{X}]] \\
&= (n-2) \left[\sigma_y^2 + \sigma_c^2 - \sigma_j^2 - \sum_{i \neq j} r_{ij}^2 \sigma_i^2 - \left(\sum_{k=1}^p \text{Cor}(\mathbf{Z}_j, \mathbf{X}_k) \boldsymbol{\beta}_k \right)^2 \right] \quad (13)
\end{aligned}$$

$$= (n-2)c_j(\sigma_y^2 + \sigma_c^2), \quad (14)$$

where $c_j = 1 - \sum_i r_{ij}^2 h_{i,b}^2 - (\sum_{k=1}^p \text{Cor}(\mathbf{Z}_j, \mathbf{X}_k) \boldsymbol{\beta}_k)^2 / (\sigma_y^2 + \sigma_c^2)$.

2.3 Single-SNP analysis with covariates fitted and \mathbf{Z} known

The sum of squares required to determine expectations when covariates are fitted corresponds to quantities required for a partial F test comparing the fit of the nested models:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (15)$$

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_j\boldsymbol{\gamma}_j + \boldsymbol{\varepsilon}, \quad (16)$$

where $\boldsymbol{\beta}$, $\boldsymbol{\gamma}_j$ are fixed and $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_e^2 \mathbf{I})$. We use subscripts 1 and 2 to denote models (15) and (16).

$$\begin{aligned} E[SSR_{j,1}|\mathbf{Z}, \mathbf{X}] &= E[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})] \\ &= \text{Tr}(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})']) \\ &= \text{Tr}(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\{\mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2\mathbf{I}_n - \frac{\sigma_e^2}{n}\mathbf{J} + \mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\}) \\ &= \text{Tr}(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\{\mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2\mathbf{I}_n + \mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\}) \\ &= \text{Tr}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X} + \sigma_e^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}) \\ &= \text{Tr}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X}) + p\sigma_e^2 + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}. \end{aligned} \quad (17)$$

$$\begin{aligned} E[SSR_{j,2}|\mathbf{Z}, \mathbf{X}] &= E\left[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})' \begin{pmatrix} \mathbf{X} & \mathbf{Z}_j \end{pmatrix} \begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_j \\ \mathbf{Z}_j'\mathbf{X} & \mathbf{Z}_j'\mathbf{Z}_j \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}' \\ \mathbf{Z}_j' \end{pmatrix} (\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})\right] \\ &= \text{Tr}\left(\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_j \\ \mathbf{Z}_j'\mathbf{X} & \mathbf{Z}_j'\mathbf{Z}_j \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}' \\ \mathbf{Z}_j' \end{pmatrix} E[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'] \begin{pmatrix} \mathbf{X} & \mathbf{Z}_j \end{pmatrix}\right) \\ &= \text{Tr}\left(\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_j \\ \mathbf{Z}_j'\mathbf{X} & \mathbf{Z}_j'\mathbf{Z}_j \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}' \\ \mathbf{Z}_j' \end{pmatrix} \{\mathbf{Z}\Sigma\mathbf{Z}' + \sigma_e^2\mathbf{I}_n + \mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\} \begin{pmatrix} \mathbf{X} & \mathbf{Z}_j \end{pmatrix}\right) \\ &= \text{Tr}\left(\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_j \\ \mathbf{Z}_j'\mathbf{X} & \mathbf{Z}_j'\mathbf{Z}_j \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X} & \mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{Z}_j \\ \mathbf{Z}_j'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X} & \mathbf{Z}_j'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{Z}_j \end{pmatrix}\right) + (p+1)\sigma_e^2 \\ &\quad + \text{Tr}\left(\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_j \\ \mathbf{Z}_j'\mathbf{X} & \mathbf{Z}_j'\mathbf{Z}_j \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}'\mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\mathbf{Z}_j \\ \mathbf{Z}_j'\mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\mathbf{X} & \mathbf{Z}_j'\mathbf{X}\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{X}'\mathbf{Z}_j \end{pmatrix}\right). \end{aligned} \quad (18)$$

To simplify, we expand the inverse in (18) using the block inversion formula, obtaining

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}_j \\ \mathbf{Z}_j'\mathbf{X} & \mathbf{Z}_j'\mathbf{Z}_j \end{pmatrix}^{-1} = \begin{pmatrix} (\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}_j'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} & -(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1} \\ -\mathbf{B}_j^{-1}\mathbf{Z}_j'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} & \mathbf{B}_j^{-1} \end{pmatrix}, \quad (19)$$

where $\mathbf{B}_j = \mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j$. Substituting (19) into (18) and noting that only the diagonal blocks of (19) are of interest, we find for the blocks associated with $\mathbf{Z}\Sigma\mathbf{Z}'$,

$$\begin{aligned} & \left((\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad -(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1} \right) \begin{pmatrix} \mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X} \\ \mathbf{Z}'_j\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X} \end{pmatrix} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X} \end{aligned} \quad (20)$$

$$\begin{pmatrix} -\mathbf{B}_j^{-1}\mathbf{Z}'_j\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} & \mathbf{B}_j^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{Z}_j \\ \mathbf{Z}'_j\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{Z}_j \end{pmatrix} = \mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{Z}_j. \quad (21)$$

For the blocks associated with $\mathbf{X}\beta\beta'\mathbf{X}'$, we get

$$\begin{aligned} & \left((\mathbf{X}'\mathbf{X})^{-1} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \quad -(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1} \right) \begin{pmatrix} \mathbf{X}'\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{X} \\ \mathbf{Z}'_j\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{X} \end{pmatrix} \\ &= \beta\beta'\mathbf{X}'\mathbf{X} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{X} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{X} = \beta\beta'\mathbf{X}'\mathbf{X} \\ & \begin{pmatrix} -\mathbf{B}_j^{-1}\mathbf{Z}'_j\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} & \mathbf{B}_j^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{X}'\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{Z}_j \\ \mathbf{Z}'_j\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{Z}_j \end{pmatrix} = -\mathbf{B}_j^{-1}\mathbf{Z}'_j\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{Z}_j + \mathbf{B}_j^{-1}\mathbf{Z}'_j\mathbf{X}\beta\beta'\mathbf{X}'\mathbf{Z}_j = \mathbf{0} \end{aligned} \quad (22)$$

Substituting (20)-(23) into (18), we obtain

$$\begin{aligned} E[SSR_{j,2}|\mathbf{Z}, \mathbf{X}] &= \text{Tr}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X}) \\ &\quad + \text{Tr}(\mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{Z}_j) + (p+1)\sigma_e^2 + \text{Tr}(\beta\beta'\mathbf{X}'\mathbf{X}) \\ &= \text{Tr}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X}) - \text{Tr}(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}') \\ &\quad + \text{Tr}(\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}') + (p+1)\sigma_e^2 + \beta'\mathbf{X}'\mathbf{X}\beta \\ &= \text{Tr}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X}) + \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}') \\ &\quad + (p+1)\sigma_e^2 + \beta'\mathbf{X}'\mathbf{X}\beta. \end{aligned} \quad (24)$$

The difference in regression sum of squares between model 1 (15) and model 2 (16) is

$$\begin{aligned} E[SSR_{j,2} - SSR_{j,1}|\mathbf{Z}, \mathbf{X}] &= \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}') + (p+1)\sigma_e^2 - p\sigma_e^2 \\ &\quad + \text{Tr}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X}) - \text{Tr}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}'\mathbf{X}) + \beta'\mathbf{X}'\mathbf{X}\beta - \beta'\mathbf{X}'\mathbf{X}\beta \\ &= \sigma_e^2 + \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}'_j(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}'), \end{aligned} \quad (25)$$

and the error sum of squares for model 2 is

$$\begin{aligned}
E[SSE_{j,2}|\mathbf{Z}, \mathbf{X}] &= E[(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})'(\mathbf{y} - \bar{\mathbf{y}}\mathbf{1})] - E[SSR_{j,2}|\mathbf{Z}, \mathbf{X}] \\
&= \text{Tr}(\mathbf{Z}\Sigma\mathbf{Z}') + (n-1)\sigma_e^2 + \text{Tr}(\mathbf{X}\beta\beta'\mathbf{X}') - E[SSR_2|\mathbf{Z}, \mathbf{X}] \\
&= \text{Tr}(\mathbf{Z}\Sigma\mathbf{Z}') + (n-1)\sigma_e^2 + \beta'\mathbf{X}'\mathbf{X}\beta - \text{Tr}(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\Sigma\mathbf{Z}') \\
&\quad - \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}') - (p+1)\sigma_e^2 - \beta'\mathbf{X}'\mathbf{X}\beta \\
&= \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}') - \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}') \\
&= (n-p-2)\sigma_e^2
\end{aligned} \tag{26}$$

2.4 Expected sums of squares when \mathbf{Z}, \mathbf{X} unknown and independent

Having marginalised expectations over \mathbf{Z} in section 2.2, the next step is to marginalise over \mathbf{X} in order to deal with the case that \mathbf{X} is not available. Since there are usually no reference panels of covariates, stronger assumptions are needed than for \mathbf{Z} . One plausible assumption is that \mathbf{X} and \mathbf{Z} are independent, so the \mathbf{X} includes covariates and not confounders. In addition, and without loss of generality, we will assume that the columns of \mathbf{X} are mutually independent.

2.4.1 Single-SNP analysis with no covariates fitted, $\text{Cor}(\mathbf{X}, \mathbf{Z}) = \mathbf{0}$.

The expected sum of squares when \mathbf{X} is independent of \mathbf{Z} and was included in the analysis model correspond to (12) and (14) with $\text{Cor}(\mathbf{Z}_j, \mathbf{X}_k) = 0 \forall k$. These expectations are,

$$E[SSR_j] = E[E[SSR_j|\mathbf{X}]] = \sigma_y^2 + \sigma_c^2 + (n-2) \left(\sigma_j^2 + \sum_{i \neq j} r_{ij}^2 \sigma_i^2 \right) \tag{27}$$

$$E[SSE_j] = E[E[SSE_j|\mathbf{X}]] = (n-2) \left(\sigma_y^2 + \sigma_c^2 - \sigma_j^2 - \sum_{i \neq j} r_{ij}^2 \sigma_{\alpha_i}^2 \right) = (n-2)c_j(\sigma_y^2 + \sigma_c^2), \tag{28}$$

where $c_j = 1 - \sum_i r_{ij}^2 h_{i,b}^2$.

2.4.2 Single-SNP analysis with covariates fitted, $\text{Cor}(\mathbf{X}, \mathbf{Z}) = \mathbf{0}$.

We note that $\mathbf{Z}\Sigma\mathbf{Z}' = \mathbf{Z}_j\Sigma_j\mathbf{Z}_j' + \sum_{i \neq j} \mathbf{Z}_i\Sigma_i\mathbf{Z}_i'$ and substitute the result into the RHS of (25), obtaining

$$\begin{aligned}
&= \sigma_e^2 + \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\Sigma\mathbf{Z}') \\
&= \sigma_e^2 + \sum_{i \neq j} \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_i\Sigma_i\mathbf{Z}_i') + \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\Sigma_j\mathbf{Z}_j')
\end{aligned} \tag{29}$$

For $\text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\boldsymbol{\Sigma}_j\mathbf{Z}_j') = \mathbf{Z}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\boldsymbol{\Sigma}_j = \mathbf{B}_j\boldsymbol{\Sigma}_j$, we need

$$\mathbb{E}[\mathbf{B}_j] = \mathbb{E}[\mathbf{Z}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j] = n-1 - \mathbb{E}[(n-1) \sum_{k=1}^p \hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k)^2] \approx n-1 - (n-1)\frac{p}{n-1} = n-1-p, \quad (30)$$

For the second trace in the expected value, rather than with \mathbf{Z}_i directly, we use the equivalent

$$\mathbf{Z}_i = \mathbf{Z}_j\hat{r}_{ij} + \hat{\mathbf{e}}_{ij},$$

where \hat{r}_{ij} is the sample correlation and the estimated coefficient for a simple linear regression of \mathbf{Z}_i on \mathbf{Z}_j , since \mathbf{Z} is standardised, and $\hat{\mathbf{e}}_{ij}$ is the estimated residual vector from the same regression.

$$\begin{aligned} & \text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_i\boldsymbol{\Sigma}_i\mathbf{Z}_i') \\ &= (\hat{r}_{ij}\mathbf{Z}_j' + \hat{\mathbf{e}}_{ij}')(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}_j'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{Z}_j\hat{r}_{ij} + \hat{\mathbf{e}}_{ij})\boldsymbol{\Sigma}_i. \end{aligned} \quad (31)$$

After expanding out terms and noting that by definition $\mathbf{e}_{ij}'\mathbf{Z}_j = 0$, (31) corresponds to

$$\hat{r}_{ij}^2\mathbf{B}_j\boldsymbol{\Sigma}_i - 2\hat{r}_{ij}\hat{\mathbf{e}}_{ij}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\boldsymbol{\Sigma}_i + \hat{\mathbf{e}}_{ij}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}_j\mathbf{B}_j^{-1}\mathbf{Z}_j'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{e}}_{ij}\boldsymbol{\Sigma}_i. \quad (32)$$

We can write (32) in sample covariance notation, which due to standardisation is equivalent to correlation except for the case of $\hat{\text{Cov}}(\hat{\mathbf{e}}_{ij}, \mathbf{X}_k) = \hat{\text{Cor}}(\hat{\mathbf{e}}_{ij}, \mathbf{X}_k)\sqrt{1-\hat{r}_{ij}^2}$. Thus we have

$$\begin{aligned} & (n-1-p)\hat{r}_{ij}^2\sigma_j^2 - 2(n-1)\hat{r}_{ij}\sqrt{1-\hat{r}_{ij}^2} \sum_{k=1}^p \hat{\text{Cor}}(\hat{\mathbf{e}}_{ij}, \mathbf{X}_k)\hat{\text{Cor}}(\mathbf{Z}_i, \mathbf{X}_k)\sigma_j^2 + \\ & (1-\hat{r}_{ij}^2)(n-1) \sum_{k=1}^p \sum_{l=1}^p \sum_{m=1}^p \sum_{o=1}^p \frac{\hat{\text{Cor}}(\hat{\mathbf{e}}_{ij}, \mathbf{X}_k)\hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k)\hat{\text{Cor}}(\hat{\mathbf{e}}_{ij}, \mathbf{X}_l)\hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_l)}{1-\hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_m)\hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_o)}\sigma_j^2. \end{aligned} \quad (33)$$

To take the expectation of (33) with respect to \mathbf{X} , note if $k = l = m = o$, we have to scale the square of a t -distributed random variable $(n-2)\hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k)^2/(1-\hat{\text{Cor}}(\mathbf{Z}_j, \mathbf{X}_k)^2)$, giving us an approximate expectation,

$$\left((n-1-p)\hat{r}_{ij}^2 + (1-\hat{r}_{ij}^2)\frac{np}{(n-2)^2} \right) \sigma_j^2, \quad (34)$$

which, marginalising over \mathbf{Z} and assuming n large, can be simplified:

$$\frac{n-p}{n-1} \left((n-2)r_{ij}^2 + 1 \right) \sigma_j^2 + \frac{p}{n-2} \left(1 - \frac{(n-2)r_{ij}^2}{n-1} - \frac{1}{n-1} \right) \sigma_j^2. \quad (35)$$

The resulting approximation for $E[SSR_{j,2} - SSR_{j,1}]$ is,

$$\begin{aligned}
& \sigma_e^2 + (n-1-p)\sigma_j^2 + \left\{ \frac{(n-1-p)(n-2) - p}{n-1} \right\} \sum_{i \neq j} r_{ij}^2 \sigma_j^2 + \frac{(n-1-p) + p}{n-1} \sum_{i \neq j} \sigma_j^2 \\
& = \sigma_e^2 + (n-1-p)\sigma_j^2 + (n-2-p) \sum_{i \neq j} r_{ij}^2 \sigma_j^2 + \sum_{i \neq j} \sigma_j^2 \\
& = \sigma_y^2 + (n-2-p) \sum_i r_{ij}^2 \sigma_j^2.
\end{aligned} \tag{36}$$

For $E[SSE_{j,2}]$, we need the expectation of $\text{Tr}((\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Z}\mathbf{\Sigma}\mathbf{Z}')$, which is,

$$\sum_{i=1}^m E[(n-1)\sigma_j^2 + (n-1) \sum_{k=1}^p \hat{\text{Cor}}(\mathbf{Z}_i, \mathbf{X}_k)^2 \sigma_j^2] \approx \sum_{i=1}^m (n-1)\sigma_j^2 + (n-1) \frac{p}{n-1} \sigma_j^2 = (n-1-p) \sum_{i=1}^m \sigma_j^2 \tag{37}$$

Combined with (26) and (36), we obtain

$$\begin{aligned}
E[SSE_{j,2}] & \approx (n-2-p)\sigma_e^2 + (n-1-p) \sum_{i=1}^m \sigma_j^2 - (n-2-p)\sigma_j^2 - (n-2-p) \sum_{i \neq j} r_{ij}^2 \sigma_j^2 - \sum_{i=1}^m \sigma_j^2 \\
& = (n-2-p) \left(\sigma_y^2 - \sigma_j^2 - \sum_{i \neq j} r_{ij}^2 \sigma_j^2 \right)
\end{aligned} \tag{38}$$

2.5 Single-SNP analysis with covariates fitted, $\text{Cor}(\mathbf{X}, \mathbf{Z}) \neq \mathbf{0}$

In a multiple regression, the t statistic for a covariate i is conditional on all other covariates fitted in the model. Hence, statistics about \mathbf{Z}_j from a regression with both \mathbf{Z}_j and \mathbf{X} fitted can be written as testing for an effect for the covariate $\hat{\mathbf{e}}_j$, where $\hat{\mathbf{e}}_j$ is the estimated residual from the regression,

$$\mathbf{Z}_j = \mathbf{X}\hat{\gamma}_j + \hat{\mathbf{e}}_j. \tag{39}$$

Since \mathbf{Z} is standardised, the estimated residual has sample variance $1 - R_j^2$, where R_j^2 is the coefficient of determination, or equivalently the inverse of the variance inflation factor [2]. Standardising $\hat{\mathbf{e}}_j$ to $\tilde{\mathbf{e}}_j$ we obtain to a linear model equivalent to (4):

$$\tilde{\mathbf{y}} = \mathbf{1}\mu + \tilde{\mathbf{e}}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\alpha} \sim \mathcal{N}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}), \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, a\sigma_e^2\mathbf{I}), \tag{40}$$

where $\tilde{\boldsymbol{\Sigma}}_{jj} = (1 - R_j^2)\sigma_j^2$, $\tilde{\mathbf{y}}$ is the covariate corrected phenotype and $a = (n-2-p)(n-2)$. We still need to convert the LD score from the observed to the residualised scale. To do this, we note

$$\begin{aligned}
\hat{r}_{ij} = \hat{\text{Cor}}(\mathbf{Z}_i, \mathbf{Z}_j) &:= \hat{\text{Cov}}(\mathbf{Z}_i, \mathbf{Z}_j) \\
&= \hat{\text{Cov}}(\mathbf{X}\hat{\gamma}_i, \mathbf{X}\hat{\gamma}_j) + \hat{\text{Cov}}(\mathbf{X}\hat{\gamma}_i, \hat{\mathbf{e}}_j) + \hat{\text{Cov}}(\hat{\mathbf{e}}_i, \mathbf{X}\hat{\gamma}_j) + \hat{\text{Cov}}(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) \\
&= \hat{\gamma}_i' \text{Var}[\mathbf{X}] \hat{\gamma}_j + \hat{\gamma}_i' \hat{\text{Cov}}(\mathbf{X}, \hat{\mathbf{e}}_j) + \hat{\text{Cov}}(\hat{\mathbf{e}}_i, \mathbf{X}) \hat{\gamma}_j + \hat{\text{Cov}}(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) \\
&= \hat{\gamma}_i' \text{Var}[\mathbf{X}] \hat{\gamma}_j + \hat{\text{Cor}}(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) \sqrt{(1-R_i^2)(1-R_j^2)},
\end{aligned} \tag{41}$$

giving an “LD score” on the residualised scale of,

$$\hat{r}'_{ij} = \text{Cor}(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) = \frac{\hat{r}_{ij} - \hat{\gamma}_i' \hat{\text{Var}}[\mathbf{X}] \hat{\gamma}_j}{\sqrt{(1-R_i^2)(1-R_j^2)}}. \tag{42}$$

substituting (40) and (42) into (8) and (10) respectively, with $\mathbf{X} = \mathbf{0}$, obtaining

$$\begin{aligned}
\text{E}[SSR_j | \mathbf{Z}, \mathbf{X}] &= a\sigma_e^2 + (n-1)(1-R_j^2)\sigma_j^2 + (n-1) \sum_{i \neq j} (\hat{r}'_{ij})^2 (1-R_i^2) \sigma_i^2 \\
&= a\sigma_e^2 + (n-1)(1-R_j^2)\sigma_j^2 + (n-1) \sum_{i \neq j} \frac{(\hat{r}_{ij} - \hat{\gamma}_i' \hat{\text{Var}}[\mathbf{X}] \hat{\gamma}_j)^2 (1-R_i^2) \sigma_i^2}{(1-R_i^2)(1-R_j^2)} \\
&= a\sigma_e^2 + (n-1)(1-R_j^2)\sigma_j^2 + (n-1) \sum_{i \neq j} \frac{(\hat{r}_{ij} - \hat{\gamma}_i' \hat{\text{Var}}[\mathbf{X}] \hat{\gamma}_j)^2}{(1-R_j^2)} \sigma_i^2.
\end{aligned} \tag{43}$$

For $SSE_j | \mathbf{Z}, \mathbf{X}$, the expectation can be determined by taking the expectations of the individual components, $SST | \mathbf{X} - SSR_j | \mathbf{Z}, \mathbf{X}$. These expected values are,

$$\begin{aligned}
\text{E}[SST | \mathbf{X}] &= (n-1) \left(a\sigma_e^2 + \sum_{i=1}^m (1-R_i^2) \sigma_i^2 \right) \\
\text{E}[SSE_j | \mathbf{Z}, \mathbf{X}] &= (n-1) \left(a\sigma_e^2 + \sum_{i \neq j} (1-R_i^2) \sigma_i^2 - \sum_{i \neq j} \frac{(\hat{r}_{ij} - \hat{\gamma}_i' \hat{\text{Var}}[\mathbf{X}] \hat{\gamma}_j)^2}{(1-R_j^2)} \sigma_i^2 \right) - a\sigma_e^2.
\end{aligned} \tag{44}$$

2.6 Meta-analysis test statistics

We have shown that the definition of h_{SNP}^2 is dependent on the covariates fitted. Now, we show how this impacts estimation of h_{SNP}^2 from meta analysis summary statistics, including when there are individuals shared between studies. We focus on two types of meta-analysis, sample size weighted [3], where $T_{j, \text{Meta}}$ is calculated as,

$$\frac{\sum_s T_{j,s} w_s}{\sqrt{\sum_s w_s^2}}, \quad (45)$$

with $w_s = \sqrt{n_s}$, the sample size of study s , and inverse variance weighted [4], where a new estimate of the effect and the variance of the estimate is obtained and weights w_s are equal to $1/\text{Var}[\hat{\alpha}_{j,s}]$,

$$\hat{\alpha}_{j,\text{Meta}} = \frac{\sum_s w_s \hat{\alpha}_{j,s}}{\sum_s w_s}, \quad \text{Var}[\hat{\alpha}_{j,\text{Meta}}] = \frac{1}{\sum_s w_s}. \quad (46)$$

Ideally, each study s in the set S would have performed identical analyses. In practice models may differ between studies. If the covariates missing from the analysis model in study s are independent of the SNPs \mathbf{Z} , then unbiasedness of the estimated SNP effect $\hat{\alpha}_{j,s}$ would still hold for all s [5] if the response is continuous [6]. This is sufficient for meta-analysis to gain statistical power at identifying causal loci without introducing bias, but insufficient for h_{SNP}^2 estimation. To prove this point, consider the case where there are always effects associated with \mathbf{X} , where \mathbf{X} is standardised and independent of \mathbf{Z} but \mathbf{y} is not standardised. Now partition S into the set S_1 where the analysis model ignored covariates and S_2 which did not,

$$\mathbf{y}_s = \mathbf{1}\mu + \mathbf{Z}_{j,s}\alpha_j + \boldsymbol{\varepsilon}, \quad \text{if } s \in S_1 \quad (47)$$

$$\mathbf{y}_s = \mathbf{1}\mu + \mathbf{Z}_{j,s}\alpha_j + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \text{if } s \in S_2. \quad (48)$$

For a sample-size weighted meta-analysis, and remembering that $\mathbf{X}\boldsymbol{\beta}$ is absorbed into $\boldsymbol{\varepsilon}_s$ for studies $s \in S_1$,

$$\begin{aligned} \text{E}[T_{j,\text{Meta}}^2] &= \frac{\text{E}_{\mathbf{Z}}[\text{E}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}}[(\sum_s T_{j,s} w_s)^2]]}{\sum_s w_s^2} \\ &= \frac{\text{E}_{\mathbf{Z}}[\text{E}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}}[(\sum_s \hat{\alpha}_{j,s} w_s / \sqrt{\text{Var}[\hat{\alpha}_{j,s}]})^2]]}{\sum_s w_s^2} \\ &\approx \frac{1}{\sum_s w_s^2} \text{E}_{\mathbf{Z}} \left[\text{E}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}} \left[\left(\sum_s \frac{\text{E}[\hat{\alpha}_{j,s} | \boldsymbol{\alpha}, \boldsymbol{\varepsilon}] w_s}{\sqrt{\text{E}[\text{Var}[\hat{\alpha}_{j,s}]}}} \right)^2 \right] \right] \\ &= \frac{1}{\sum_s w_s^2} \text{E}_{\mathbf{Z}} \left[\text{E}_{\boldsymbol{\alpha},\boldsymbol{\varepsilon}} \left[\left(\sum_s \frac{(\sum_i \hat{r}_{ij,s} \boldsymbol{\alpha}_i + c_s (\mathbf{Z}'_{j,s} \mathbf{Z}_{j,s})^{-1} \mathbf{Z}'_j \boldsymbol{\varepsilon}_s) w_s}{\sqrt{\text{E}[\text{Var}[\hat{\alpha}_{j,s}]}}} \right)^2 \right] \right] \\ &= \frac{1}{\sum_s w_s^2} \left(\text{E}_{\mathbf{Z}} \left[\sum_{s,s'} \frac{\text{E}_{\boldsymbol{\alpha}}[(\sum_i \hat{r}_{ij,s} \boldsymbol{\alpha}_i)(\sum_i \hat{r}_{ij,s'} \boldsymbol{\alpha}_i)] w_s w_{s'}}{\sqrt{\text{E}[\text{Var}[\hat{\alpha}_{j,s}]] \text{E}[\text{Var}[\hat{\alpha}_{j,s'}]]}} \right] + \text{E}_{\boldsymbol{\varepsilon}} \left[\left(\sum_s \frac{c_s (\mathbf{Z}'_{j,s} \mathbf{Z}_{j,s})^{-1} \mathbf{Z}'_j \boldsymbol{\varepsilon}_s w_s}{\sqrt{\text{E}[\text{Var}[\hat{\alpha}_{j,s}]}}} \right)^2 \right] \right) \end{aligned}$$

$$= \frac{1}{\sum_s w_s^2} \left(\sum_i \sigma_i^2 \sum_{s,s'} \frac{\hat{r}_{ij,s} \hat{r}_{ij,s'} w_s w_{s'}}{\sqrt{E[\text{Var}[\hat{\alpha}_{j,s}]] E[\text{Var}[\hat{\alpha}_{j,s'}]]}} + \sum_{s,s'} \frac{(\sigma_e^2 + \mathbb{1}_{s,s' \in S_1} \sigma_c^2) c_s c_{s'} w_s w_{s'} (n_{ss'} - 1)}{(n_s - 1)(n_{s'} - 1) \sqrt{E[\text{Var}[\hat{\alpha}_{j,s}]] E[\text{Var}[\hat{\alpha}_{j,s'}]]}} \right) \quad (49)$$

where $c_s = 1$ and $c_s = \sqrt{(n_s - 1)/(n_s - p - 1)}$ for analysis models (47) and (48). To complete the derivation of the approximate expectation, we need

$$E[\text{Var}[\hat{\alpha}_{j,s}]] = \begin{cases} \frac{\sigma_y^2 + \sigma_c^2 - \sum_i r_{ij}^2 \sigma_i^2}{n_s - 1} \approx \frac{\sigma_y^2 + \sigma_c^2}{n_s - 1} \approx \frac{\sigma_y^2 + \sigma_c^2}{w_s^2} & \text{if } c_s = 1 \\ \frac{\sigma_y^2 - \sum_i r_{ij}^2 \sigma_i^2}{n_s - p - 1} \approx \frac{\sigma_y^2}{n_s - p - 1} \approx \frac{\sigma_y^2}{c_s^2 w_s^2} & \text{if } c_s = \sqrt{(n_s - p - 1)/(n_s - 1)}. \end{cases} \quad (50)$$

If $c_s = 1 \forall s$ or $c_s = \sqrt{(n_s - p - 1)/(n_s - 1)} \forall s$ and assume that $p \ll n_s \forall s \Rightarrow c_s \approx 1$, (49) can be reduced to,

$$\begin{aligned} E[T_{j,\text{Meta}}^2] &= \frac{1}{\sum_s w_s^2} \left(\sum_i \sigma_i^2 \sum_{s,s'} \frac{\hat{r}_{ij,s} \hat{r}_{ij,s'} c_s c_{s'} w_s^2 w_{s'}^2}{\sigma_T^2} + (\sigma_T^2 - \sum_i \sigma_i^2) \sum_{s,s'} \frac{c_s^2 c_{s'}^2 w_s^2 w_{s'}^2 (n_{ss'} - 1)}{(n_s - 1)(n_{s'} - 1) \sigma_T^2} \right), \\ &\approx \frac{1}{\sum_s w_s^2} \left(\sum_i \sigma_i^2 \sum_{s,s'} \frac{\hat{r}_{ij,s} \hat{r}_{ij,s'} w_s^2 w_{s'}^2}{\sigma_T^2} + (\sigma_T^2 - \sum_i \sigma_i^2) \sum_{s,s'} \frac{n_{ss'}}{\sigma_T^2} \right), \end{aligned} \quad (51)$$

where $\sigma_T^2 = \sigma_y^2 + \sigma_c^2$ and noting that $\sigma_e^2 + \mathbb{1}_{s,s' \in S_1} \sigma_c^2 = \sigma_T^2 - \sum_i \sigma_i^2$ when no covariate correction was applied, and σ_y^2 if covariate correction was applied. When the approximate expectation for the sample LD in studies s and s' ,

$$E[\hat{r}_{ij,s} \hat{r}_{ij,s'}] = r_{ij}^2 + \frac{n_{ss'}}{n_s n_{s'}} \quad (52)$$

is substituted into (51), we find

$$\begin{aligned} E[T_{j,\text{Meta}}^2] &\approx \frac{1}{\sum_s w_s^2} \left(\sum_i \frac{r_{ij}^2 \sigma_i^2}{\sigma_T^2} \sum_s w_s^2 \sum_{s'} w_{s'}^2 + \frac{\sigma_T^2 - \sum_i \sigma_i^2 + \sum_i \sigma_i^2}{\sigma_T^2} \sum_{s,s'} n_{ss'} \right), \\ &= \sum_s n_s \sum_i \frac{r_{ij}^2 \sigma_i^2}{\sigma_T^2} + \frac{\sum_{s,s'} n_{ss'}}{\sum_s n_s}, \\ &\approx \begin{cases} 1 + \frac{2}{n} \sum_{s \neq s'} n_{ss'} + n \sum_i r_{ij}^2 h_{i,b}^2 & \text{if } c_s = 1 \quad \forall s \\ 1 + \frac{2}{n} \sum_{s \neq s'} n_{ss'} + n \sum_i r_{ij}^2 h_{i,a}^2 & \text{if } c_s = \sqrt{(n_s - 1)/(n_s - p - 1)} \quad \forall s \end{cases}. \end{aligned} \quad (53)$$

Unlike the single study considered earlier, $E[T_{j,\text{Meta}}^2]$ can be inflated in such a way that the inclusion of an intercept term is sufficient to obtain unbiased estimates of h_{SNP}^2 . However this inflation is due to individuals being included in multiple studies, not the explicit ignoring of confounders we focus on in this paper. If studies from sets S_1 and S_2 , are combined, a meaningful interpretation of the

slope as a function of a heritability parameter is no longer possible, as in

$$\begin{aligned}
E[T_{j,\text{Meta}}^2] &= \frac{\sum_i \sigma_i^2}{\sum_s w_s^2} \left(\sum_{s,s' \in S_1} \frac{\sqrt{(n_s-1)(n_{s'}-1)} \hat{r}_{ij,s} \hat{r}_{ij,s'} w_s w_{s'}}{\sigma_y^2 + \sigma_c^2} + \sum_{s,s' \in S_2} \frac{\sqrt{(n_s-p-1)(n_{s'}-p-1)} \hat{r}_{ij,s} \hat{r}_{ij,s'} w_s w_{s'}}{\sigma_y^2} \right. \\
&\quad \left. + \sum_{s \in S_1, s' \in S_2} \frac{\sqrt{(n_s-1)(n_{s'}-p-1)} w_s w_{s'} \hat{r}_{ij,s} \hat{r}_{ij,s'}}{\sigma_y^2 \sqrt{1 + \sigma_c^2/\sigma_y^2}} \right) \\
&\quad + \frac{\sigma_e^2}{\sum_s w_s^2} \left(\sum_{s,s' \in S_2} \frac{w_s w_{s'}}{\sigma_y^2} + \sum_{s \in S_1, s' \in S_2} \frac{w_s w_{s'}}{\sigma_y^2 \sqrt{1 + \sigma_c^2/\sigma_y^2}} \right) + \frac{\sigma_e^2 + \sigma_c^2}{\sum_s w_s^2} \sum_{s,s' \in S_1} \frac{w_s w_{s'}}{\sigma_y^2 + \sigma_c^2}, \quad (54)
\end{aligned}$$

the denominator changes depending on the study's choice of S_1 or S_2 . Similar problems arise in the case of inverse variance weighting, both for the test statistic $T_{j,\text{Meta}}^2$ and the effect size $\hat{\alpha}_{j,\text{Meta}}$. The expected value of $\hat{\alpha}_{j,\text{Meta}}^2$ from an inverse variance weighted meta-analysis is,

$$\begin{aligned}
E[\hat{\alpha}_{j,\text{Meta}}^2] &= E\left[\left(\frac{\sum_s \hat{\alpha}_{j,s} w_s}{\sum_s w_s}\right)^2\right] \approx \frac{\sum_{s,s'} E[(\hat{\alpha}_{j,s} \hat{\alpha}_{j,s'}] w_s w_{s'}}{(\sum_s w_s)^2} \\
&= \sum_i \sigma_i^2 \frac{\sum_{s,s'} (r_{ij}^2 + \frac{n_{ss'}}{n_s n_{s'}}) w_s w_{s'}}{(\sum_s w_s)^2} + \sum_{s,s'} (\sigma_e^2 + \mathbb{1}_{s,s' \in S_2} \sigma_c^2) \frac{n_{ss'} c_s c_{s'} w_s w_{s'}}{(n_s-1)(n_{s'}-1)(\sum_s w_s)^2} \\
&\approx \sum_i r_{ij}^2 \sigma_i^2 \frac{(\sum_s w_s)^2}{(\sum_s w_s)^2} + \sum_{s,s'} (\sigma_e^2 + \mathbb{1}_{s,s' \in S_2} \sigma_c^2) \frac{n_{ss'} c_s c_{s'} w_s w_{s'}}{n_s n_{s'} (\sum_s w_s)^2} + \sum_i \sigma_i^2 \sum_{s,s'} \frac{n_{ss'} w_s w_{s'}}{n_s n_{s'} (\sum_s w_s)^2} \\
&= \sum_i r_{ij}^2 \sigma_i^2 + \sum_{s,s'} (\sigma_e^2 + \mathbb{1}_{s,s' \in S_1} \sigma_c^2) \frac{n_{ss'} c_s c_{s'} w_s w_{s'}}{n_s n_{s'} (\sum_s w_s)^2} + \sum_i \sigma_i^2 \sum_{s,s'} \frac{n_{ss'} w_s w_{s'}}{n_s n_{s'} (\sum_s w_s)^2}, \quad (55)
\end{aligned}$$

The expected values of the weights in (55) are,

$$w_s \approx 1/E[\text{Var}[\hat{\alpha}_{j,s}]] = \frac{n_s-1}{\sigma_y^2 + \sigma_c^2 - \sum_i r_{ij}^2 \sigma_i^2} \approx \frac{n_s-1}{\sigma_y^2 + \sigma_c^2}, \quad \text{if } c_s = 1 \quad (56)$$

$$w_s \approx 1/E[\text{Var}[\hat{\alpha}_{j,s}]] = \frac{n_s-p-1}{\sigma_y^2 - \sum_i r_{ij}^2 \sigma_i^2} \approx \frac{n_s-p-1}{\sigma_y^2} = \frac{c_s^2(n_s-1)}{\sigma_y^2} \quad \text{if } c_s = \sqrt{(n_s-p-1)/(n_s-1)} \quad (57)$$

which upon substitution into (55) and again assuming $p \ll n$ so that $c_s \approx 1 \forall s$ gives,

$$E[\hat{\alpha}_{j,\text{Meta}}^2] \approx \begin{cases} \left(\frac{1}{\sum_s n_s-1} + \frac{2 \sum_{s \neq s'} n_{ss'}}{(\sum_s n_s-1)^2} \right) (\sigma_y^2 + \sigma_c^2) + \sum_i r_{ij}^2 \sigma_i^2 & \text{if } c_s = 1 \quad \forall s \\ \left(\frac{1}{\sum_s n_s-1} + \frac{2 \sum_{s \neq s'} n_{ss'}}{(\sum_s n_s-1)^2} \right) \sigma_y^2 + \sum_i r_{ij}^2 \sigma_i^2 & \text{if } c_s = \sqrt{\frac{n_s-1}{n_s-p-1}} \quad \forall s \\ \frac{(\sigma_y^2 + \sigma_c^2) \sum_{s,s' \in S_1} n_{ss'}}{(\sum_s n_s-1)^2} + \frac{\sigma_y^2 \sum_{s,s' \in S_2} n_{ss'}}{(\sum_s n_s-1)^2} + \frac{\sigma_y^2 \sum_{s \in S_1, s' \in S_2} n_{ss'}}{(\sum_s n_s-1)^2} + \sum_i r_{ij}^2 \sigma_i^2 & \text{in general.} \end{cases} \quad (58)$$

Similarly, the test statistic $T_{j,\text{Meta}}^2 = \hat{\alpha}_{j,\text{Meta}}^2 \sum_s w_s$ has an approximate expected value of

$$E[T_{j,\text{Meta}}^2] \approx \begin{cases} 1 + \frac{2 \sum_{s \neq s'} n_{ss'}}{\sum_s n_s} + \sum_s n_s \sum_i r_{ij}^2 h_{i,b}^2 & \text{if } c_s = 1 \forall s \\ 1 + \frac{2 \sum_{s \neq s'} n_{ss'}}{\sum_s n_s - 1} + \sum_s n_s \sum_i r_{ij}^2 h_{i,a}^2 & \text{if } c_s = \sqrt{\frac{n_s - 1}{n_s - p - 1}} \forall s \\ \left(\frac{(\sigma_y^2 + \sigma_c^2) \sum_{s,s' \in S_1} n_{ss'} + \sigma_y^2 (\sum_{s,s' \in S_2} n_{ss'} + \sum_{s \in S_1, s' \in S_2} n_{ss'})}{(\sum_s n_s)^2} + \sum_i r_{ij}^2 \sigma_i^2 \right) \left(\sum_{s \in S_1} \frac{n_s - 1}{\sigma_y^2 + \sigma_c^2} + \sum_{s \in S_2} \frac{n_s - p - 1}{\sigma_y^2} \right) & \text{in general.} \end{cases} \quad (59)$$

so like sample weighted meta-analysis shows additive inflation in the presence of shared individuals, and cannot estimate a meaningful heritability parameter when different studies did not apply equivalent covariate corrections. The presence of additive inflation should be viewed as analogous to the expectation of the cross product of test statistics given in Bulik-Sullivan et al. [7] in order to estimate genetic correlation.

In the main text we showed that if $\text{Cov}[\mathbf{X}, \mathbf{Z}] = 0$, the SumHer model intercept is 1 and there is no need to fit A/C . In Figure S9, we show what happens if we repeat this analysis with sample overlap. Specifically the meta-analysis now consists of two studies, one with all 8000 individuals and the with 4000 of the individuals. Based on the theory given in section 2.6, either h_{SNPa}^2 or h_{SNPb}^2 can still be estimated without/with covariate correction in both studies, but with model intercept $2 \times 4000 / 12000 = 5/3$. Our simulations are consistent with this theory under both LDAK and GCTA models.

2.7 Mixed model test statistics

For the purposes of this section, we will assume the phenotype model

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\alpha} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_e^2 \mathbf{I}), \quad (60)$$

where μ is an intercept, \mathbf{Z} a $n \times m$ matrix of standardised SNP genotypes, $\boldsymbol{\alpha}$ a vector of SNP effect sizes, and $\boldsymbol{\Sigma}$ a diagonal matrix with j th entry σ_j^2 . We assume the model fitted in the association study is,

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{Z}_j \boldsymbol{\alpha}_j + \mathbf{Z} \boldsymbol{\alpha}_{-j} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\alpha}_{-j} \sim \mathcal{N}(\mathbf{0}, \tilde{\boldsymbol{\Sigma}}), \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \tilde{\sigma}_e^2 \mathbf{I}), \quad (61)$$

where the effect of SNP j is treated as a fixed effect, and the elements of $\tilde{\boldsymbol{\Sigma}}$ and $\tilde{\sigma}_e^2$ need not match the equivalent entries in the true model. The test statistic for testing the hypothesis whether $\boldsymbol{\alpha}_j = 0$ calculated under the assumed model of (61) is,

$$T_{j,LMM}^2 = \mathbf{y}' \mathbf{V}^{-1} \mathbf{Z}_j (\mathbf{Z}_j' \mathbf{V}^{-1} \mathbf{Z}_j)^{-1} \mathbf{Z}_j' \mathbf{V}^{-1} \mathbf{y} \quad (62)$$

where $\mathbf{V} = \mathbf{Z}_{-j}\tilde{\Sigma}\mathbf{Z}_{-j}' + \sigma_e^2\mathbf{I}$. Note that part of the simplicity of (62) is due to the centring required to standardise \mathbf{Z} , which implies $\text{cov}(\hat{\mu}, \hat{\alpha}_j) = 0$. Now

$$\begin{aligned}
\mathbb{E}[T_{j, LMM}^2] &= \mathbb{E}[\mathbf{y}'\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{y}] \\
&= \text{Tr}(\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{V}^{-1}\mathbb{E}[\mathbf{y}\mathbf{y}']) \\
&= \text{Tr}(\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{V}^{-1}(\mu^2\mathbf{1}\mathbf{1}' + \sigma_j^2\mathbf{Z}_j\mathbf{Z}_j' + \mathbf{Z}_{-j}\Sigma_{-j}\mathbf{Z}_{-j}' + \sigma_e^2\mathbf{I})) \\
&= \text{Tr}(\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{V}^{-1}(\sigma_j^2\mathbf{Z}_j\mathbf{Z}_j' + \mathbf{Z}_{-j}\Sigma_{-j}\mathbf{Z}_{-j}' + \sigma_e^2\mathbf{I})) \\
&= \mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j\sigma_j^2 + \text{Tr}(\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{V}^{-1}(\mathbf{Z}_{-j}\Sigma_{-j}\mathbf{Z}_{-j}' + \sigma_e^2\mathbf{I})) \\
&= \mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j\sigma_j^2 + \text{Tr}(\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{V}^{-1}(\mathbf{Z}_{-j}\Sigma_{-j}\mathbf{Z}_{-j}' + \sigma_e^2\mathbf{I})). \tag{63}
\end{aligned}$$

Lastly, define $\mathbf{V}_{\text{True}} = \mathbf{Z}_{-j}\Sigma_{-j}\mathbf{Z}_{-j}' + \sigma_e^2\mathbf{I}$ and $\mathbf{E} = \mathbf{V}_{\text{True}} - \mathbf{V}$. Then we can write (63) as,

$$\begin{aligned}
\mathbb{E}[T_{j, LMM}^2] &= \mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j\sigma_j^2 + \text{Tr}(\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}\mathbf{Z}_j'\mathbf{V}^{-1}(\mathbf{V} + \mathbf{E})) \\
&= \mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j\sigma_j^2 + 1 + \mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{E}\mathbf{V}^{-1}\mathbf{Z}_j(\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j)^{-1}. \tag{64}
\end{aligned}$$

For the purposes of trying to re-estimate h_{SNP}^2 from mixed model association statistics, while $\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_j = \text{Var}[\hat{\beta}]$ is a readily obtainable quantity, $\mathbf{Z}_j'\mathbf{V}^{-1}$ is not. More importantly, re-estimation of h_{SNP}^2 would only be of interest in the case $\mathbf{E} \neq \mathbf{0}$.

3 Additional simulation results

3.1 Phenotypes generated assuming under a GCTA model

To show that our conclusions are not dependent on the simulation model, we simulated phenotypes using a GCTA model and analysed using LDSC. Results are presented in Figures S1-S4.

In Figure S1(a-c), we see the same patterns as in the LDAK simulations, namely ignoring covariates means only h_{SNPb}^2 can be estimated, while conditioning on covariates, means only h_{SNPa}^2 can be estimated from T_j^2 , while both h_{SNPa}^2 and h_{SNPb}^2 can be estimated from $n\hat{\alpha}_j^2$. By changing the subset of SNPs, the degree of negative bias when estimating h_{SNP}^2 from $n\hat{\alpha}_j^2$ and fixing the intercept to one when covariates were adjusted for is less than for the LDAK simulations presented in the main paper.

For meta-analysis test statistics (Figure S2), we find as expected the same patterns as in the LDAK simulations. If there is consistency in the GWAS analysis method, a meaningful definition of h_{SNP}^2 is estimated. If there is inconsistency in analysis models between component GWAS s , the estimate of h_{SNP}^2 obtained from a meta-analysis is not interpretable.

In Figure S3, we see that only full adjustment for confounders in the original GWAS results

in an unbiased estimate of h_{SNPa}^2 . As with the LDAK simulations, biased estimates of h_{SNP}^2 were obtained with both the case of no adjustment or partial adjustment at the GWAS stage. Compared to the LDAK simulations, the level of bias in h_{SNP}^2 was lower when simulating under a GCTA model (Figures S3 and S4).

Like the LDAK simulations, we find that the positive bias in the estimate of h_{SNPb}^2 and the scale of A/C (Figure S4) for $C1$, $C2$, and $C3$ phenotypes correspond to the proportion of phenotypic variation attributed to confounding. Again like the LDAK simulations, and as expected both from the positive linear relationship between LD score and a_j and from results in [8], the estimates of A/C were below the mean level of confounding.

3.2 Estimation of h_{SNP}^2 when applying filtering

In many of the simulations with unaccounted confounding, we generate S_j with values sufficiently high that SNP j would be removed from analysis using the filtering methods proposed in [9] as an outlier. We applied this filtering approach to see what impact it had on the estimate of h_{SNP}^2 and A/C from phenotypes simulated under both a GCTA and LDAK model.

The results presented for LDAK simulations in Figures S5-S6 are equivalent to Figures 3-4 in the main paper, but with filtering applied before summary statistic h_{SNP}^2 analysis. The filtering consisted of removing SNPs where $S_j > 80$ and also those SNPs in linkage disequilibrium (defined as $r_{ij}^2 > 10$) with such SNPs from analysis. Similarly the results presented for GCTA simulations in Figures S7-S8 are equivalent to Figures S3-S4 in the main paper, but with filtering applied before summary statistic h_{SNP}^2 analysis.

While as expected, the estimates of h_{SNP}^2 drop compared to the no filtering case for both LDAK and GCTA simulations, in no case was the drop sufficient to eliminate bias. While some of the GCTA simulations may appear to be unbiased, with the average estimate of h_{SNP}^2 sitting on the black line, it is always the red line that is the estimable definition of h_{SNP}^2 in these simulations. Furthermore, as the filtering is based on a fixed cut-off, as the proportion of phenotypic variation, $\sigma_c^2/(\sigma_y^2 + \sigma_c^2)$ due to confounding increased, an increasing number of SNPs were removed in the filtering process. This meant the patterns linking $\sigma_c^2/(\sigma_y^2 + \sigma_c^2)$ to the level of bias in the estimate of h_{SNP}^2 are broken as seen in Figures S6 and S8. From a statistical perspective, this indicates ad-hoc filtering approaches may be insufficient to remove all influential SNPs that could bias the estimation of parameters.

4 Supporting Figures

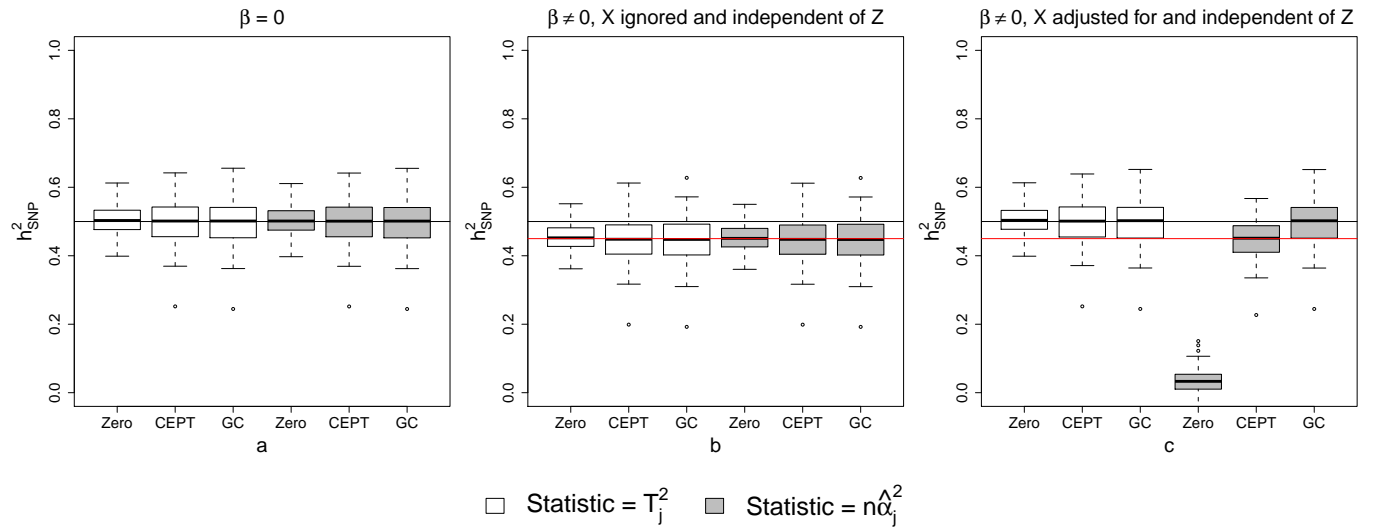


Figure S1: Estimates of h^2_{SNP} obtained from LDSC analysis of GWAS summary statistics constructed for phenotypes simulated under a GCTA model. The black and red horizontal lines indicate the values of h^2_{SNPa} and h^2_{SNPb} . Zero, CEPT and GC refer to no, A and C confounding terms in the analysis model. (a) Phenotypes with no covariate effects. (b) Phenotypes with covariate effects but \mathbf{X} ignored in the analysis. (c) Phenotypes with covariate effects and \mathbf{X} adjusted for in the analysis.

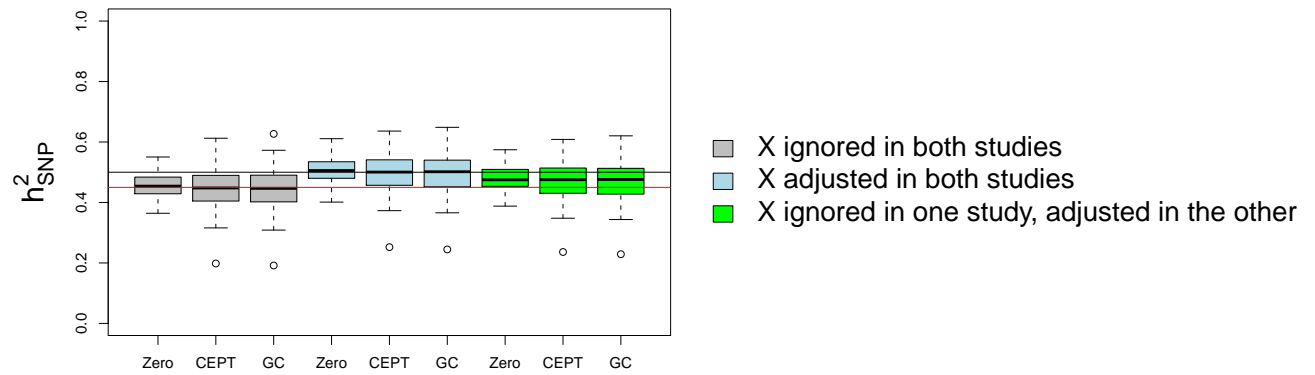


Figure S2: Estimates of h^2_{SNP} obtained from LDSC analysis of summary statistics calculated from a meta-analysis of two GWAS constructed from phenotypes simulated under a GCTA model.

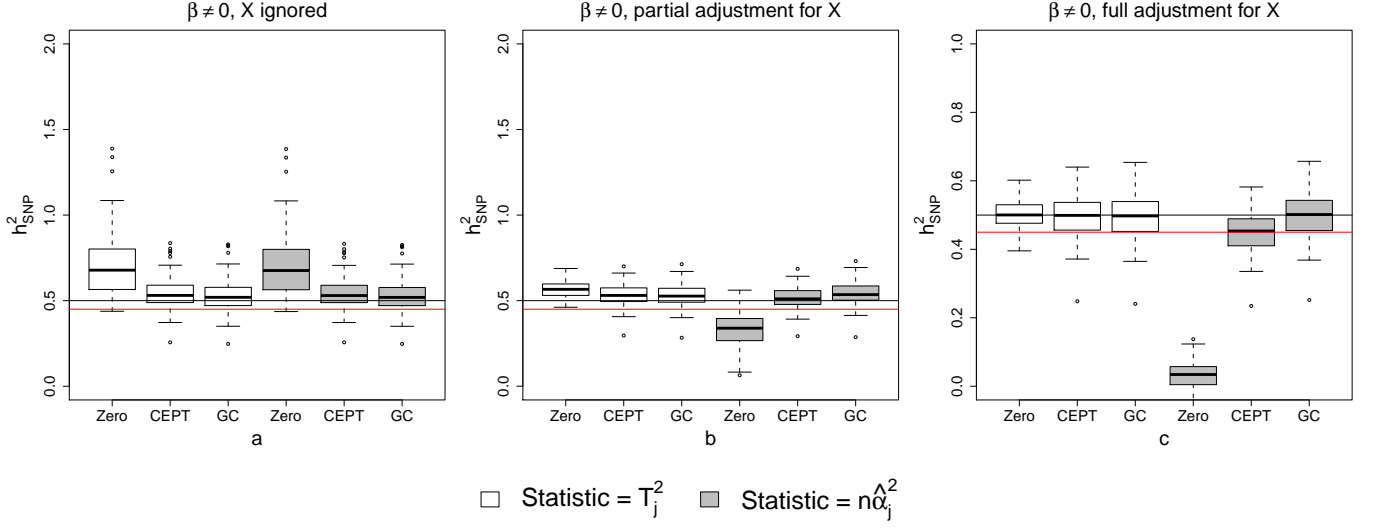


Figure S3: Similar to Figure S1, but here GWAS phenotypes (specifically C1) are subject to confounding: phenotype means differ among three subpopulations that each consist of three sub-subpopulations. The sub-populations were the same as used in the LDAK simulations. Estimates of h_{SNP}^2 from a GWAS with (a) no covariate adjustment, (b) adjustment for the three subpopulations but not the sub-subpopulations, (c) full covariate adjustment. Note that the y-axis differs among (a), (b) and (c).

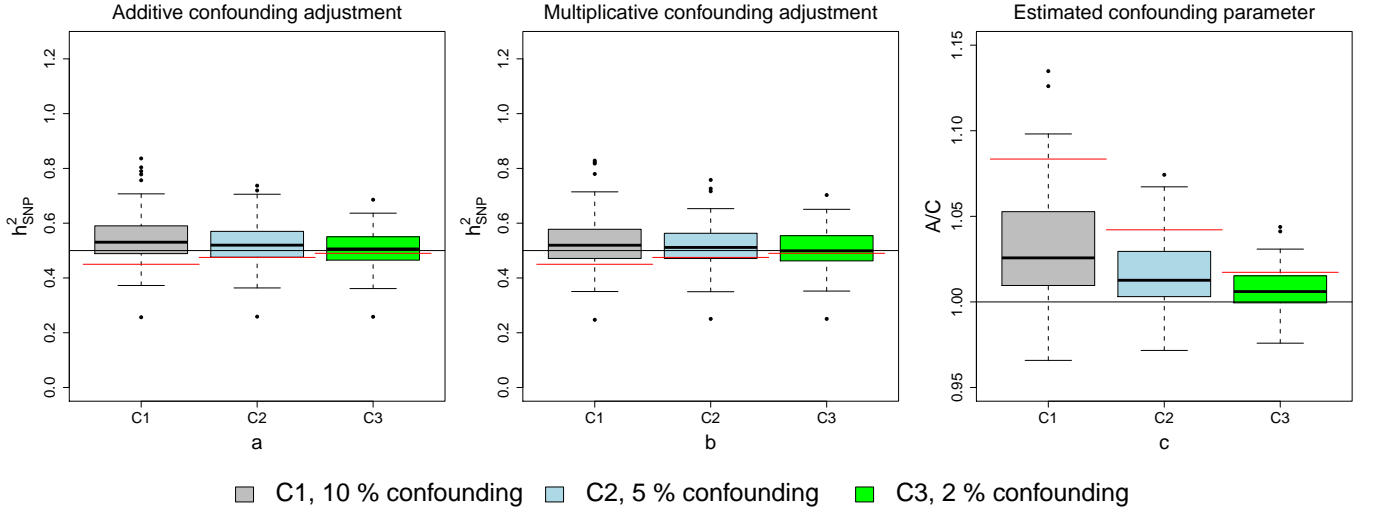


Figure S4: Estimating h_{SNP}^2 and confounding parameters from phenotypes with differing proportion of phenotypic variation due to confounding when $h_{\text{SNP}_a}^2 = 0.5$. The confounding corresponds to ignoring sub-populations, which are the same as used in the LDAK simulations. The black lines in (a,b) indicates the simulated value of $h_{\text{SNP}_a}^2$ and the red lines the simulated value of $h_{\text{SNP}_b}^2$. In (c), the black line corresponds to $A/C = 1$, which is normally assumed to indicate no confounding and the red line the mean level of confounding, estimated as $\bar{S}_j - 1 - n \sum_i r_{ij}^2 h_{j,b}^2$. Note that the y-axis differs between plots.

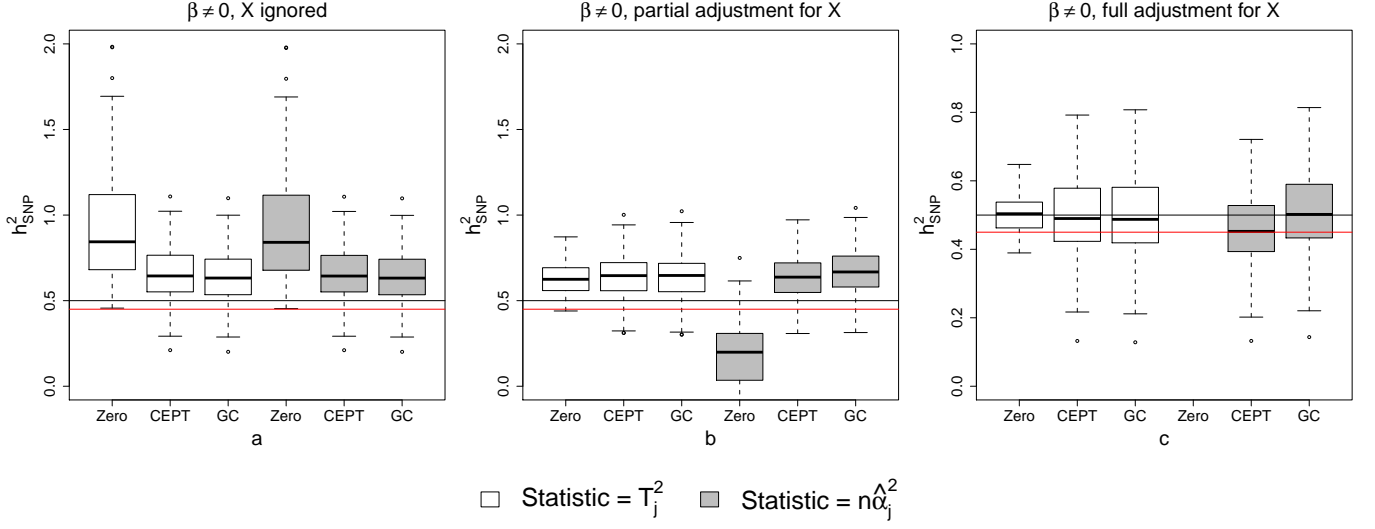


Figure S5: Same as Figures 3 in the main paper, but with filtering out of SNPs where $S_j > 80$ prior to summary statistic h_{SNP}^2 analysis. Estimates of h_{SNP}^2 from a GWAS with (a) no covariate adjustment, (b) adjustment for the three subpopulations but not the sub-subpopulations, (c) full covariate adjustment. Note that the y -axis differs between plots.

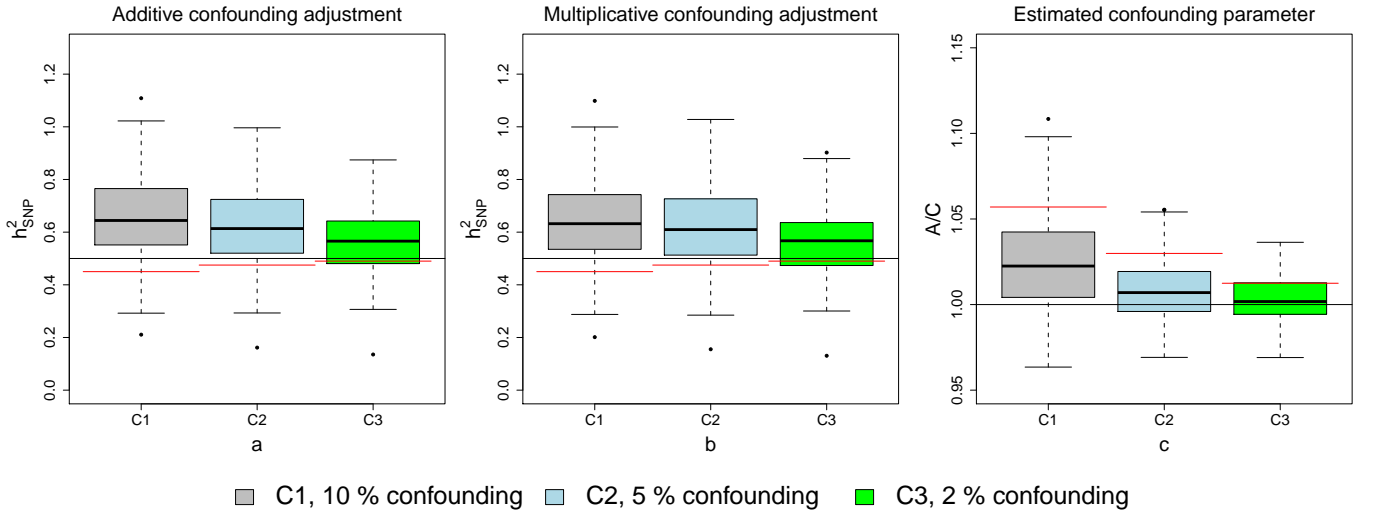


Figure S6: Same as Figure 4 in the main paper, but with filtering out of SNPs where $S_j > 80$ prior to summary statistic h_{SNP}^2 analysis. The black lines in (a,b) indicates the simulated value of h_{SNPa}^2 and the red lines the simulated value of h_{SNPb}^2 . In (c), the black line corresponds to $A/C = 1$, which is normally assumed to indicate no confounding and the red line the mean level of confounding, estimated as $\bar{S}_j - 1 - n \sum_i r_{ij}^2 h_{j,b}^2$. Note that the y -axis differs between plots.

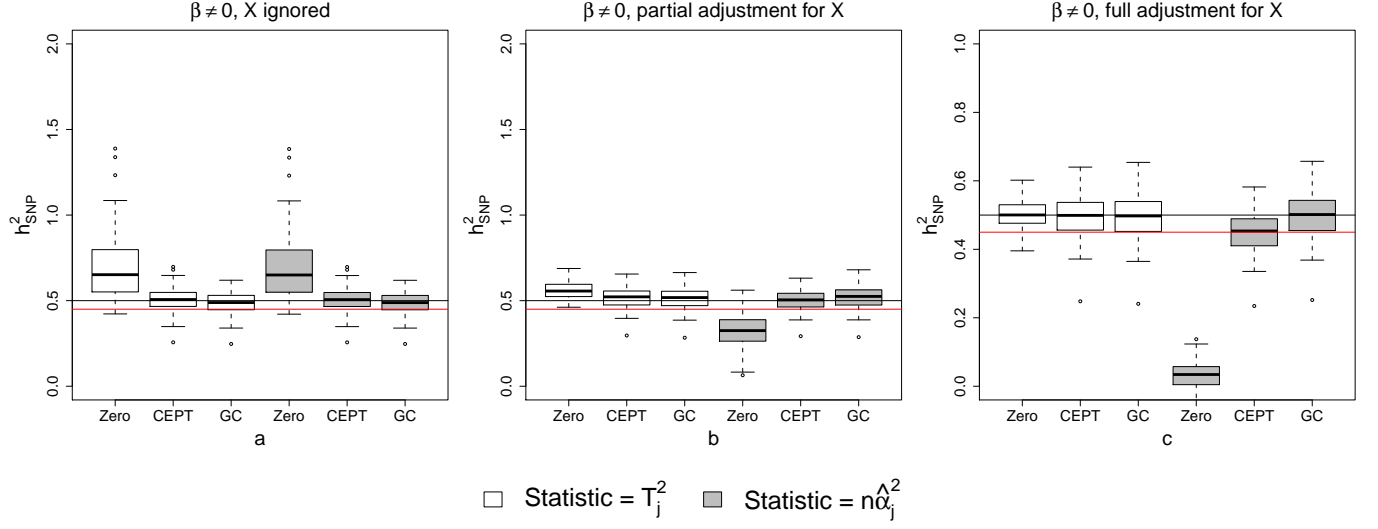


Figure S7: Same as Figure S3, but with filtering out of SNPs where $S_j > 80$ prior to summary statistic h_{SNP}^2 analysis. Estimates of h_{SNP}^2 from a GWAS with (a) no covariate adjustment, (b) adjustment for the three subpopulations but not the subsubpopulations, (c) full covariate adjustment. Note that the y -axis differs among (a), (b) and (c).

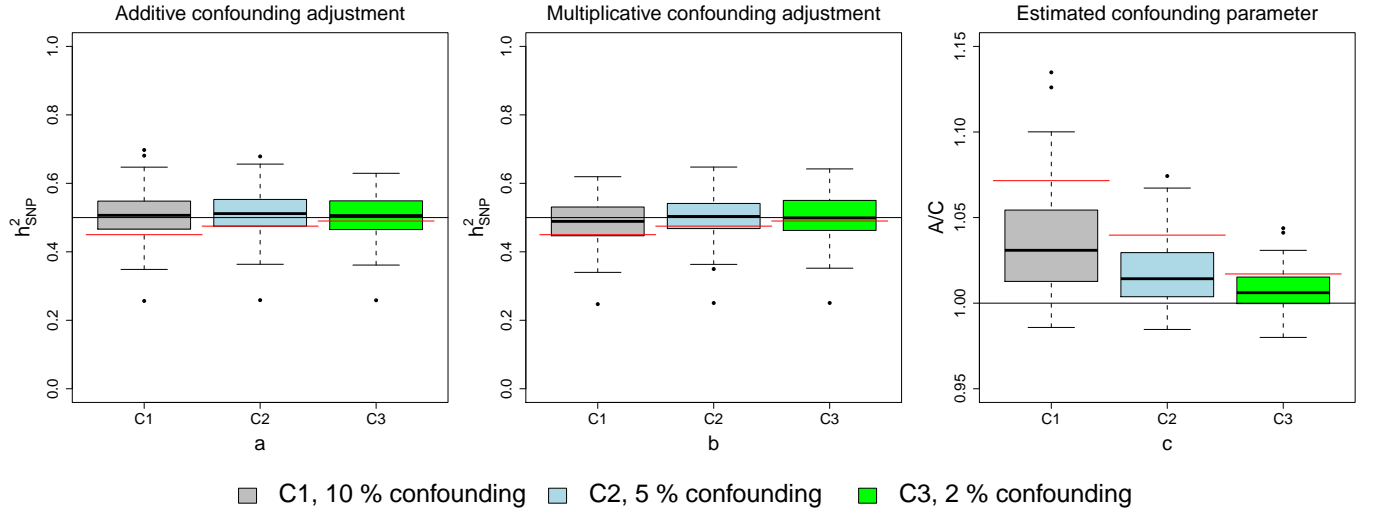


Figure S8: Same as Figure S4, but with filtering out of SNPs where $S_j > 80$ prior to summary statistic h_{SNP}^2 analysis. The black lines in (a,b) indicates the simulated value of $h_{\text{SNP}_a}^2$ and the red lines the simulated value of $h_{\text{SNP}_b}^2$. In (c), the black line corresponds to $A/C = 1$, which is normally assumed to indicate no confounding and the red line the mean level of confounding, estimated as $\bar{S}_j - 1 - n \overline{\sum_i r_{ij}^2 h_{j,b}^2}$. Note that the y -axis differs between (a,b) and (c).

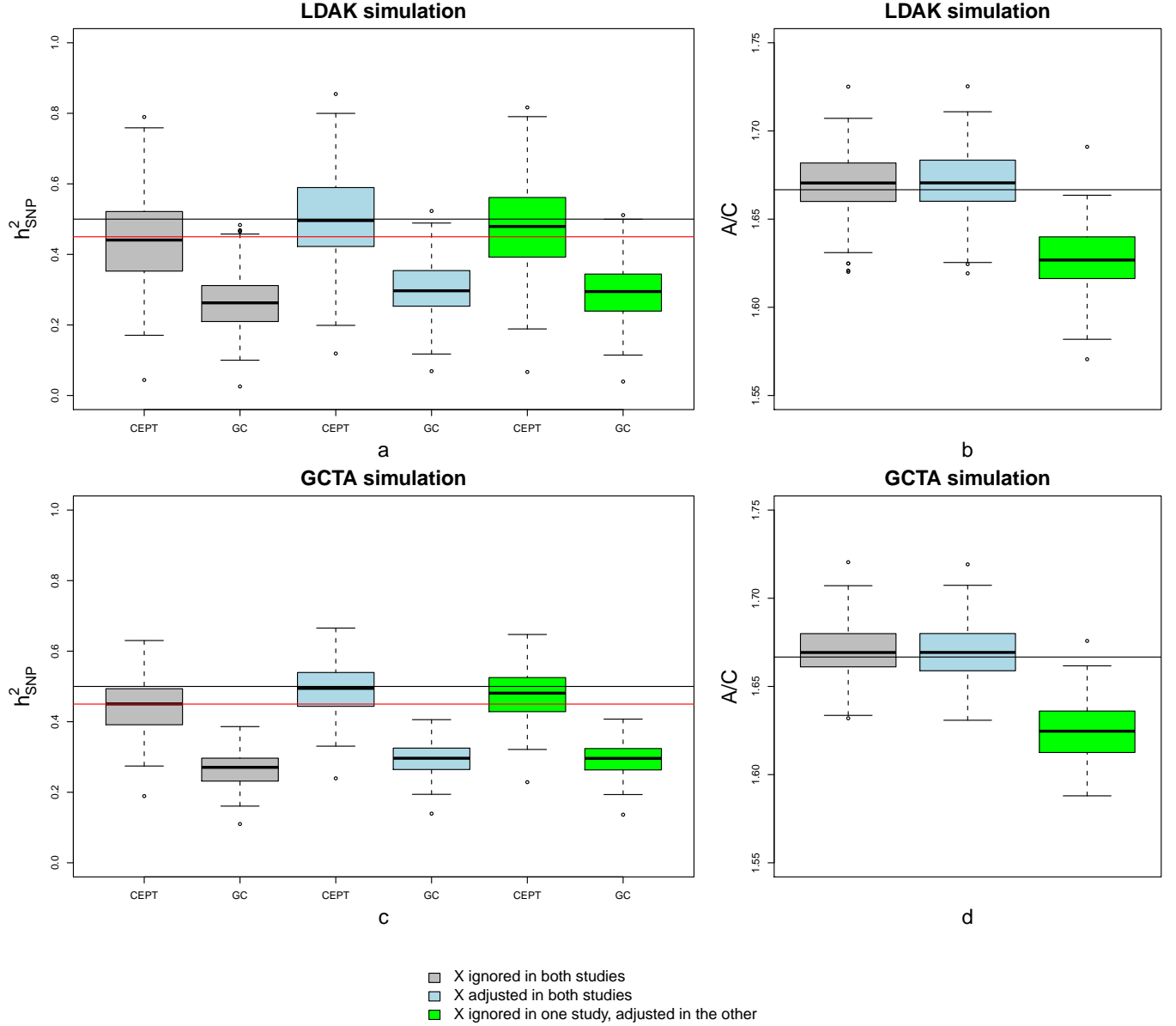


Figure S9: Similar as Figure 2 in the main paper, but now the meta-analysis includes overlapping 4000 individuals out of a total of 12,000. The black and red horizontal lines in (a,c) indicate the values of h^2_{SNPa} and h^2_{SNPb} . The black horizontal line in (b,d) indicates the expected intercept. CEPT and GC refer to A and C confounding terms in the analysis. Plots (a) and (c) gives estimates of h^2_{SNP} obtained under either an A or C view of the confounding parameter, while plots (b) and (d) gives estimates of the confounding parameter. Plots (a,b) are obtained from phenotypes simulated from a LD simulation, (c,d) from phenotypes simulated from a GCTA model.

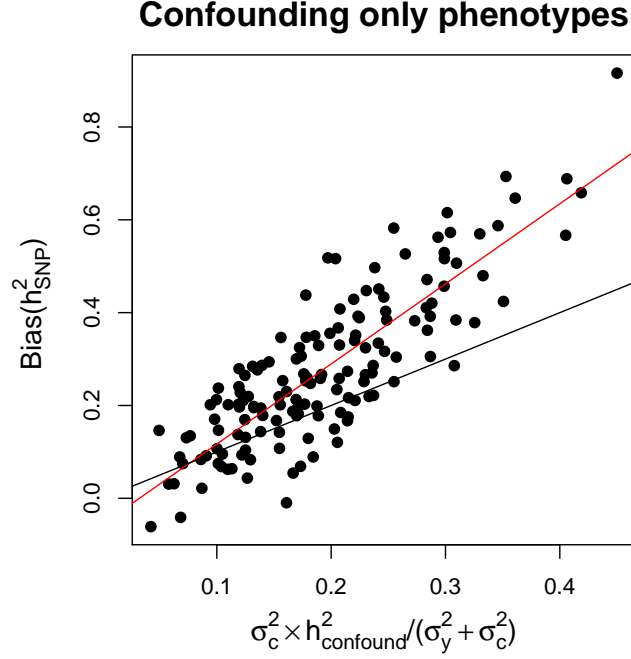


Figure S10: Estimates of h^2_{SNP} (called h^2_{confound}) when the phenotype is the confounding effect $\mathbf{X}\beta$ only, scaled to match the contribution of confounding to a C1 phenotype compared to the observed bias in h^2_{SNP} when confounding correction was not and was included in the GWAS. The $y = x$ line is shown in black, which corresponds to the bias in h^2_{SNP} being completely explained by the confounding-only test statistic. The red line is the fit of simple linear regression of the bias on scaled h^2_{confound} . Bias was calculated for each phenotype using $\hat{h}^2_{\text{SNP,ignored}} - \hat{h}^2_{\text{SNP,corrected}}(h^2_{\text{SNPb}}/h^2_{\text{SNPa}})$, i.e by comparing estimates from a confounding ignored and a confounding corrected GWAS while accounting for the change in estimable h^2_{SNP} parameter. One confounding-only phenotypes did not produce convergent estimates so was removed for comparison purposes.

References

- [1] Brendan K. Bulik-Sullivan, Po-Ru Loh, Hilary Finucane, Stephan Ripke, Jian Yang, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Nick Patterson, Mark J. Daly, Alkes L. Price, and Benjamin M. Neale. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.*, 47:291–295, 2015.
- [2] Henri Theil. *Principles of Econometrics*. John Wiley & Sons, 1971.
- [3] S. Stouffer, L. DeVinney, and E. Suchmen. *The American soldier, Volume 1: Adjustment during army life*. 1949.
- [4] Cristen J. Willer, Yun Li, and Goncalo R. Abecasis. Metal: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, 26(17):2190–2191, 2010.
- [5] O. Kempthorne. Why randomize. *J. Stat. Plan. Inference*, 1:1–25, 1977.
- [6] M.H. Gail, S. Wieand, and S. Piantadosi. Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates. *Biometrika*, 71(3):431–444, 1984.
- [7] B. K. Bulik-Sullivan, H. K. Finucane, V. Anttila, A. Gusev, F. R. Day, P.-R. Loh, ReproGen Consortium, Psychiatric Genomics Consortium, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium, Laramie Duncan, J. R. B. Perry, Nick Patterson, Elise B. Robinson, M. J. Daly, A. L. Price, and B. M. Neale. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.*, 47(11):1236–1241, 2015.
- [8] Ronald DeVlaming, Magnus Johannesson, Patrik K.E. Magnusson, M. Arfan Ikram, and Peter M. Visscher. Equivalence of LD-Score Regression and Individual-Level-Data Methods. 2017.
- [9] Jie Zheng, A. Mezut Erzurumluoglu, Benjamin L. Elsworth, John P. Kemp, Laurence Howe, Philip C. Haycock, Gibran Hemani, Katherine Tansey, Charles Laurin, Early Genetics and Life-course Epidemiology (EAGLE) Consortium, Beate St Pourcain, Nicole M. Warrington, Hilary K. Finucane, Alkes L. Price, B. K. Bulik-Sullivan, Verner Anttila, Lavinia Paternoster, Tom R. Gaunt, David M. Evans, and Benjamin M. Neale. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics*, 33(2):272–279, 2017.