Web-based Supplementary Materials for Test for Rare Variants by Environment Interactions in Sequencing Association Studies

by Xinyi Lin, Seunggeun Lee, Michael C. Wu, Chaolong Wang, Han Chen, Zilin Li and Xihong Lin

May 2015

Contents

Web Appendix 1. Bias Analysis of Burden Tests

1.1. Bias analysis of β^* under $\boldsymbol{G} - E$ independence for linear regression

We assume the true model is:

$$\mathcal{E}\left(Y_i|E_i, \boldsymbol{G}_i\right) = \alpha_1 + \alpha_2 E_i + \sum_{j=1}^{p} G_{ij}\alpha_{3j} + \sum_{j=1}^{p} G_{ij}E_i\beta_j \text{ and } \text{Var}\left(Y_i|E_i, \boldsymbol{G}_i\right) = \sigma^2. \quad (1\text{-}1)$$

The burden test linear regression model is:

$$\mathcal{E}\left(Y_i|E_i, G_i^*\right) = \alpha_1^* + \alpha_2^* E_i + \alpha_3^* G_i^* + \beta^* G_i^* E_i \text{ and } \text{Var}\left(Y_i|E_i, G_i^*\right) = \sigma_*^2. \quad (1\text{-}2)$$

Based on the true model (1-1), we can calculate the true $\mathcal{E}\left(Y_i|E_i, G_i^*\right)$ as:

$$\mathcal{E}\left(Y_i|E_i, G_i^*\right) = \alpha_1 + \alpha_2 E_i + \sum_{j=1}^{p} \mathcal{E}\left(G_{ij}|E_i, G_i^*\right)\alpha_{3j} + \sum_{j=1}^{p} \mathcal{E}\left(G_{ij}|E_i, G_i^*\right)E_i\beta_j. \quad (1\text{-}3)$$

If $\boldsymbol{G}$ and E are independent, we have:

$$\begin{aligned}\mathcal{E}\left(Y_i|E_i, G_i^*\right) &= \alpha_1 + \alpha_2 E_i + \sum_{j=1}^{p} \mathcal{E}\left(G_{ij}|G_i^*\right)\alpha_{3j} + \sum_{j=1}^{p} \mathcal{E}\left(G_{ij}|G_i^*\right)E_i\beta_j \\ &\approx \alpha_1 + \alpha_2 E_i + \left(\sum_{j=1}^{p} a_j\alpha_{3j}\right) G_i^* + \left(\sum_{j=1}^{p} a_j\beta_j\right) G_i^* E_i, \end{aligned} \quad (1\text{-}4)$$

where $a_j = \frac{\text{MAF}_j}{\sum_{k=1}^{p}\text{MAF}_k}$ and Equation (1-4) is obtained by approximating the distribution of $\left(G_{i1}, \cdots, G_{ip}|G_i^*\right)$ with Multinomial $\left(a_1, \cdots, a_p, G_i^*\right)$, giving $\mathcal{E}\left(G_{ij}|G_i^*\right) = a_j G_i^*$. This expectation $\mathcal{E}\left(G_{ij}|G_i^*\right) = a_j G_i^*$ is exact when (a) $G_{ij} \sim \text{Binom}\left(2, \text{MAF}_j\right)$, and G_{ij}'s are independent for $j = 1, \cdots, p$ and (b) $\text{MAF}_1 = \cdots = \text{MAF}_p$. It is also exact when (a) and (c) $G_i^* = 1$ holds.

1.2. Bias analysis of β^* under $\boldsymbol{G}-E$ independence for logistic regression

Assume that a disease is rare. The true model is:

$$\mathcal{E}\left(Y_i|E_i,\boldsymbol{G}_i\right)\approx \mathrm{Var}\left(Y_i|E_i,\boldsymbol{G}_i\right)\approx \exp\left[\alpha_1+\alpha_2E_i+\sum_{j=1}^{p}G_{ij}\alpha_{3j}+\sum_{j=1}^{p}G_{ij}E_i\beta_j\right]. \tag{1-5}$$

The burden test logistic regression model is:

$$\mathcal{E}\left(Y_i|E_i,G_i^*\right)\approx \mathrm{Var}\left(Y_i|E_i,G_i^*\right)\approx \exp\left[\alpha_1^*+\alpha_2^*E_i+\alpha_3^*G_i^*+\beta^*G_i^*E_i\right]. \tag{1-6}$$

Based on the true model (1-5), we can calculate the true $\mathcal{E}\left(Y_i|E_i,G_i^*\right)$ as:

$$\begin{aligned}\mathcal{E}\left(Y_i|E_i,G_i^*\right) &\approx \mathcal{E}\left[\exp\left(\alpha_1+\alpha_2E_i+\sum_{j=1}^{p}G_{ij}\alpha_{3j}+\sum_{j=1}^{p}G_{ij}E_i\beta_j\right)|E_i,G_i^*\right]\\ &\approx \exp\left(\alpha_1+\alpha_2E_i\right)\left[1+\mathcal{E}\left(\sum_{j=1}^{p}G_{ij}\alpha_{3j}+\sum_{j=1}^{p}G_{ij}E_i\beta_j|E_i,G_i^*\right)\right]\\ &\approx \exp\left(\alpha_1+\alpha_2E_i+\sum_{j=1}^{p}\mathcal{E}\left(G_{ij}|E_i,G_i^*\right)\alpha_{3j}+\sum_{j=1}^{p}\mathcal{E}\left(G_{ij}|E_i,G_i^*\right)E_i\beta_j\right)\end{aligned} \tag{1-7}$$

where we assume $\mathcal{E}\left[\left(\sum_{j=1}^{p}G_{ij}\alpha_{3j}+\sum_{j=1}^{p}G_{ij}E_i\beta_j\right)^l|E_i,G_i^*\right]$ and $\left[\mathcal{E}\left(\sum_{j=1}^{p}G_{ij}\alpha_{3j}+\sum_{j=1}^{p}G_{ij}E_i\beta_j|E_i,G_i^*\right)\right]^l$ are negligible for $l\geq 2$.
If $\boldsymbol{G}$ and E are independent, we have:

$$\begin{aligned}\mathcal{E}\left(Y_i|E_i,G_i^*\right) &= \exp\left(\alpha_1+\alpha_2E_i+\sum_{j=1}^{p}\mathcal{E}\left(G_{ij}|G_i^*\right)\alpha_{3j}+\sum_{j=1}^{p}\mathcal{E}\left(G_{ij}|G_i^*\right)E_i\beta_j\right)\\ &\approx \exp\left(\alpha_1+\alpha_2E_i+\left(\sum_{j=1}^{p}a_j\alpha_{3j}\right)G_i^*+\left(\sum_{j=1}^{p}a_j\beta_j\right)G_i^*E_i\right),\end{aligned}$$

where as discussed in the case of linear regression, the distribution of $\left(G_{i1},\cdots,G_{ip}|G_i^*\right)$ is approximated by Multinomial$\left(a_1,\cdots,a_p,G_i^*\right)$, giving $\mathcal{E}\left(G_{ij}|G_i^*\right)=a_jG_i^*$.

1.3. Bias analysis of β^* under $\boldsymbol{G} - E$ dependence

For a continuous outcome, we assume that the true model is (1-1) and the burden test model is (1-2); For a binary outcome (rare disease), we assume that the true model is (1-5) and the burden test model is (1-6). In general, the bias of β^* in (1-2) and (1-6) depends on the underlying $\boldsymbol{G} - E$ structure. In this section, we consider the special case of $\boldsymbol{G} - E$ dependence where E is binary such that:

$$(G_{i1}, \cdots, G_{ip} | G_i^*, E_i = 0) \sim \text{Multinomial}\,(b_1, \cdots, b_p, G_i^*) \text{ approximately,}$$
$$(G_{i1}, \cdots, G_{ip} | G_i^*, E_i = 1) \sim \text{Multinomial}\,(c_1, \cdots, c_p, G_i^*) \text{ approximately,}$$

where $b_j = \frac{\text{MAF}_{j0}}{\sum_{k=1}^{p} \text{MAF}_{k0}}$ and $c_j = \frac{\text{MAF}_{j1}}{\sum_{k=1}^{p} \text{MAF}_{k1}}$ for $j = 1, \cdots, p$. Similar to the calculations under $\boldsymbol{G} - E$ independence (Web Appendix 1.1. and 1.2.), we use $\mathcal{E}\,(G_{ij} | G_i^*, E_i = 0) = b_j G_i^*$ and this expectation is exact when (a) $G_{ij} | E_i = 0 \sim \text{Binom}\,(2, \text{MAF}_{j0})$, and G_{ij}'s are independent given $E_i = 0$ for $j = 1, \cdots, p$ and (b) $\text{MAF}_{10} = \cdots = \text{MAF}_{p0}$. It is also exact when (a) and (c) $G_i^* = 1$ holds. Similar conditions are required for $\mathcal{E}\,(G_{ij} | G_i^*, E_i = 1) = c_j G_i^*$.

If $\boldsymbol{G}$ and E are dependent, for linear regression, (1-3) becomes:

$$\text{(1-3)} = \mathcal{E}\,(Y_i | E_i, G_i^*) \approx \alpha_1 + \alpha_2 E_i + \left(\sum_{j=1}^{p} b_j \alpha_{3j} \right) G_i^* + \left(\sum_{j=1}^{p} c_j \alpha_{3j} - b_j \alpha_{3j} + c_j \beta_j \right) G_i^* E_i, \tag{1-8}$$

and for logistic regression with rare disease, (1-7) becomes:

$$\text{(1-7)} = \mathcal{E}\,(Y_i | E_i, G_i^*) \approx \exp\left(\alpha_1 + \alpha_2 E_i + \left(\sum_{j=1}^{p} b_j \alpha_{3j} \right) G_i^* + \left(\sum_{j=1}^{p} c_j \alpha_{3j} - b_j \alpha_{3j} + c_j \beta_j \right) G_i^* E_i \right). \tag{1-9}$$

Comparing (1-8) and the burden test model (1-2) for linear regression, and (1-9) and the burden test model (1-6) for logistic regression, we can express the parameters in the misspecified burden test models (1-2) and (1-6) in terms of the parameters in the true models (1-1) and (1-5) as:

$$\alpha_1^* = \alpha_1, \quad \alpha_2^* = \alpha_2, \quad \alpha_3^* = \sum_{j=1}^{p} b_j \alpha_{3j}, \quad \beta^* = \sum_{j=1}^{p} c_j \alpha_{3j} - b_j \alpha_{3j} + c_k \beta_j.$$

When the null hypothesis holds (i.e. $\beta_1 = \cdots = \beta_p = 0$), β^* becomes:

$$\beta^* = \sum_{j=1}^{p} (c_j - b_j)\, \alpha_{3j}.$$

β^* is thus a function of the main effects $\{\alpha_{3j}\}_{j=1}^{p}$ and is generally biased since it is capturing the main effects under the null hypothesis of no interaction effects. The bias generally worsens with increasing main effect sizes $|\alpha_{3j}|$. To understand how the strength of $\boldsymbol{G} - E$

dependence can affect the bias, we note that when $\boldsymbol{G}$ and E are independent, $c_j - b_j = 0$ for all $j = 1, \cdots, p$. In this way, $\{|c_j - b_j|\}_{j=1}^{p}$ is a measure of the $\boldsymbol{G} - E$ dependence and we would generally expect the bias of β^* to worsen with increasing $\boldsymbol{G} - E$ dependence. We note that when the main effects α_{3j} are equal, i.e. $\alpha_{3j} = \alpha_3$ for every j, the burden test will be unbiased.

Web Appendix 2. iSKAT

2.1. Estimation of the Null Model

The null model (Equation (7) in the main manuscript) is fitted using weighted ridge regression, since fitting the null model using regular maximum likelihood can give unstable results and/or the model will not converge, as the genotypes $\boldsymbol{G}$ contain rare variants that have low frequency. Define an additional weight matrix $\boldsymbol{W}_2 = \text{diag}(w_{21}, \cdots, w_{2p})$. The weighted penalized log-likelihood for fitting the null model is:

$$\ell_P(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha}) - \frac{1}{2}\lambda \boldsymbol{\alpha}_3^\intercal \boldsymbol{W}_2^{-1} \boldsymbol{W}_2^{-1} \boldsymbol{\alpha}_3,$$

where $\ell(\boldsymbol{\alpha}) = \sum_{i=1}^{n} \log\{f(Y_i|\boldsymbol{X}_i, E_i, \boldsymbol{G}_i)\}$ is the log-likelihood of Y_i under the null model, and λ is a ridge tuning parameter. The ridge parameter λ controls the goodness of fit and the complexity of the null model. When $\lambda = 0$, $\widehat{\boldsymbol{\alpha}}^{\lambda=0}$ is the usual maximum likelihood estimator. When $\lambda = \infty$, $\widehat{\boldsymbol{\alpha}}_3 = \boldsymbol{0}$ and is equivalent to assuming that there are no main effects for the rare variants. For computational efficiency, we choose the ridge parameter λ using generalized cross-validation (O'Sullivan *and others*, 1986), where we grid search λ within the range $[0.01, \sqrt{n}/\log(n)]$. We note that when $\rho = 0$, weights $w_{1j} = 1$ and $w_{2j} = 1/\sqrt{\text{var}(G_j)}$, the test statistic Q_ρ (Equation (6) in the main manuscript) reduces to the GESAT test statistic (Lin *and others*, 2013).

2.2. Derivation of the Asymptotic Null Distribution of Q_ρ for fixed ρ

Let $\hat{\lambda}$ be the estimated ridge parameter under the null model. Let $\boldsymbol{\alpha}_0$ be the true value of $\boldsymbol{\alpha}$, $\boldsymbol{\Sigma} = \text{diag}\{\text{Var}(Y_i)\}$ and $\boldsymbol{\Delta}_0 = \text{diag}\{g^{'}(\mu_i)\}$ under the null hypothesis. Define $\boldsymbol{y} = \widetilde{\boldsymbol{X}}\boldsymbol{\alpha}_0 + \boldsymbol{\Delta}_0\left(\boldsymbol{Y} - \boldsymbol{\mu}\left(\boldsymbol{\alpha}_0\right)\right)$ and $\boldsymbol{u} = \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\Delta}_0^{-1}\left(\boldsymbol{y} - \widetilde{\boldsymbol{X}}\boldsymbol{\alpha}_0\right)$. Note that each entry of $\boldsymbol{u}$ has mean zero and unit variance under the null hypothesis. Define $\boldsymbol{Z} = \Sigma^{\frac{1}{2}}\Delta_0\left(\boldsymbol{I} - \boldsymbol{H}^{\widehat{\lambda}}\right)^{\intercal}\Delta_0^{-1}\boldsymbol{S}\boldsymbol{W}_1$, where $\boldsymbol{H}^{\widehat{\lambda}} = \widetilde{\boldsymbol{X}}\left(\widetilde{\boldsymbol{X}}^{\intercal}\boldsymbol{\Delta}_0^{-1}\widetilde{\boldsymbol{X}} + \widehat{\lambda}\boldsymbol{I}_2\right)^{-1}\widetilde{\boldsymbol{X}}^{\intercal}\boldsymbol{\Delta}_0^{-1}$, and $\boldsymbol{I}_2$ is a $(q+1+p)\times(q+1+p)$ block diagonal matrix with the top $(q+1)\times(q+1)$ block diagonal matrix being 0 and the bottom $p\times p$ block diagonal matrix being $\boldsymbol{W}_2^{-1}\boldsymbol{W}_2^{-1}$. Let $\bar{\boldsymbol{Z}} = \left(\bar{Z}_1, \cdots, \bar{Z}_n\right)^{\intercal}$ where $\bar{Z}_i = \sum_{j=1}^{p} Z_{ij}/p$ and Z_{ij} is the i^{th} row and j^{th} column entry of $\boldsymbol{Z}$. Let $\boldsymbol{Z}_j$ be the jth column of $\boldsymbol{Z}$. Define $\boldsymbol{M} = \bar{\boldsymbol{Z}}(\bar{\boldsymbol{Z}}^{\intercal}\bar{\boldsymbol{Z}})^{-1}\bar{\boldsymbol{Z}}^{\intercal}$.

When $\widehat{\lambda} = o(\sqrt{n})$, we have:

$$\begin{aligned}\frac{1}{\sqrt{n}}\boldsymbol{Z}^{\intercal}\boldsymbol{u} &= \frac{1}{\sqrt{n}}\boldsymbol{W}_1\boldsymbol{S}^{\intercal}\Delta_0^{-1}\left(\boldsymbol{I} - \frac{1}{n}\widetilde{\boldsymbol{X}}\left(\frac{1}{n}\widetilde{\boldsymbol{X}}^{\intercal}\Delta_0^{-1}\widetilde{\boldsymbol{X}} + \frac{\widehat{\lambda}}{n}\boldsymbol{I}_2\right)^{-1}\widetilde{\boldsymbol{X}}^{\intercal}\Delta_0^{-1}\right)\left(\boldsymbol{y} - \widetilde{\boldsymbol{X}}\boldsymbol{\alpha}_0\right) \\ &= \frac{1}{\sqrt{n}}\boldsymbol{Z}_0^{\intercal}\boldsymbol{u} + o_p(1) \qquad (2\text{-}10)\end{aligned}$$

and

$$\begin{aligned}\frac{1}{n}Q_\rho &= \frac{1}{n}\left(\boldsymbol{y} - \widetilde{\boldsymbol{X}}\boldsymbol{\alpha}_0\right)^{\intercal}\left(\boldsymbol{I} - \boldsymbol{H}^{\widehat{\lambda}}\right)^{\intercal}\Delta_0^{-1}\boldsymbol{S}\boldsymbol{W}_1\boldsymbol{R}_\rho\boldsymbol{W}_1\boldsymbol{S}^{\intercal}\Delta_0^{-1}\left(\boldsymbol{I} - \boldsymbol{H}^{\widehat{\lambda}}\right)\left(\boldsymbol{y} - \widetilde{\boldsymbol{X}}\boldsymbol{\alpha}_0\right) + o_p(1) \\ &= \frac{1}{n}\boldsymbol{u}^{\intercal}\boldsymbol{Z}\boldsymbol{R}_\rho\boldsymbol{Z}^{\intercal}\boldsymbol{u} + o_p(1) = \frac{1}{n}\boldsymbol{u}^{\intercal}\boldsymbol{Z}_0\boldsymbol{R}_\rho\boldsymbol{Z}_0^{\intercal}\boldsymbol{u} + o_p(1),\end{aligned}$$

where $\boldsymbol{Z}_0 = \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Delta}_0\left(\boldsymbol{I} - \boldsymbol{H}_0\right)^{\intercal}\boldsymbol{\Delta}_0^{-1}\boldsymbol{S}\boldsymbol{W}_1$, and $\boldsymbol{H}_0 = \widetilde{\boldsymbol{X}}\left(\widetilde{\boldsymbol{X}}^{\intercal}\Delta_0^{-1}\widetilde{\boldsymbol{X}}\right)^{-1}\widetilde{\boldsymbol{X}}^{\intercal}\boldsymbol{\Delta}_0^{-1}$. Thus for a fixed ρ, Q_ρ asymptotically follows a mixture of chi-squares distribution $\sum d_k\chi^2_{1,k}$, where the $\chi^2_{1,k}$ are independent variables that follow a χ^2_1 distribution and d_k is the k^{th} non-zero eigenvalue of $\boldsymbol{Z}_0\boldsymbol{R}_\rho\boldsymbol{Z}_0^{\intercal}$. For finite samples, we calculate the d_k's by replacing $\boldsymbol{H}_0$ in $\boldsymbol{Z}_0$ by $\boldsymbol{H}^{\widehat{\lambda}}$. A p-value for Q_ρ is computed by inverting the characteristic function of $\sum d_k\chi^2_{1,k}$ using the Davies method (Davies, 1980).

2.3. Obtaining a p-value for Q_{iSKAT}

Calculation of the distribution of Q_{iSKAT} is challenged by the fact that the p_ρ are correlated for different ρ's. To obtain a p-value for Q_{iSKAT}, we need the joint distribution of Q_ρ for different ρ's for $0 \leq \rho \leq 1$. In this section, we first re-express the distribution of Q_ρ and show that Q_ρ can be decomposed into a weighted sum of two approximately independent random variables. We then show how this decomposition can then be used to obtain a p-value for iSKAT using the algorithm described in Lee *and others* (2012).

Let $\bar{\boldsymbol{Z}}_0 = \left(\bar{Z}_1^0, \cdots, \bar{Z}_n^0\right)^\intercal$, where $\bar{Z}_i^0 = \sum_{j=1}^p Z_{ij}^0/p$ and Z_{ij}^0 is the i^{th} row and j^{th} column entry of $\boldsymbol{Z}_0$. Let $\boldsymbol{Z}_j^0$ be the jth column of $\boldsymbol{Z}_0$. Define $\boldsymbol{M}_0 = \bar{\boldsymbol{Z}}_0(\bar{\boldsymbol{Z}}_0^\intercal \bar{\boldsymbol{Z}}_{\boldsymbol{0}})^{-1}\bar{\boldsymbol{Z}}_0^\intercal$. Since $\widehat{\lambda} = o(\sqrt{n})$,

$$\begin{aligned}
\frac{1}{n}\boldsymbol{Z}^\intercal\boldsymbol{Z} &= \frac{1}{n}\boldsymbol{W}_1\boldsymbol{S}^\intercal\Delta_0^{-1}\left(\boldsymbol{I} - \frac{1}{n}\widetilde{\boldsymbol{X}}\left(\frac{1}{n}\widetilde{\boldsymbol{X}}^\intercal\Delta_0^{-1}\widetilde{\boldsymbol{X}} + \frac{\widehat{\lambda}}{n}\boldsymbol{I}_2\right)^{-1}\widetilde{\boldsymbol{X}}^\intercal\Delta_0^{-1}\right)\Delta_0\Sigma^{1/2} \\
&\times \Sigma^{1/2}\Delta_0\left(\boldsymbol{I} - \frac{1}{n}\Delta_0^{-1}\widetilde{\boldsymbol{X}}\left(\frac{1}{n}\widetilde{\boldsymbol{X}}^\intercal\Delta_0^{-1}\widetilde{\boldsymbol{X}} + \frac{\widehat{\lambda}}{n}\boldsymbol{I}_2\right)^{-1}\widetilde{\boldsymbol{X}}^\intercal\right)\Delta_0^{-1}\boldsymbol{S}\boldsymbol{W}_1 \\
&= \frac{1}{n}\boldsymbol{Z}_0^\intercal\boldsymbol{Z}_0 + o_p(1). \qquad (2\text{-}11)
\end{aligned}$$

Using (2-10) and (2-11), it can be shown that:

$$\frac{1}{\sqrt{n}}\boldsymbol{Z}^\intercal\boldsymbol{M}\boldsymbol{u} = \frac{1}{\sqrt{n}}\boldsymbol{Z}_0^\intercal\boldsymbol{M}_0\boldsymbol{u} + o_p(1) \quad \text{and} \qquad (2\text{-}12)$$

$$\frac{1}{\sqrt{n}}\boldsymbol{Z}^\intercal\left(\boldsymbol{I} - \boldsymbol{M}\right)\boldsymbol{u} = \frac{1}{\sqrt{n}}\boldsymbol{Z}_0^\intercal\left(\boldsymbol{I} - \boldsymbol{M}_0\right)\boldsymbol{u} + o_p(1). \qquad (2\text{-}13)$$

Using (2-10), (2-12) and (2-13), it follows that:

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\left(\boldsymbol{u}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{Z}, \boldsymbol{u}^\intercal\boldsymbol{M}\boldsymbol{Z}, \boldsymbol{u}^\intercal\bar{\boldsymbol{Z}}\right)^\intercal \\
&= \frac{1}{\sqrt{n}}\left(\boldsymbol{u}^\intercal(\boldsymbol{I} - \boldsymbol{M}_0)\boldsymbol{Z}_0, \boldsymbol{u}^\intercal\boldsymbol{M}_0\boldsymbol{Z}_0, \boldsymbol{u}^\intercal\bar{\boldsymbol{Z}}_0\right)^\intercal + o_p(1) \\
&= \frac{1}{\sqrt{n}}\left(\boldsymbol{I} - \boldsymbol{1}\left(\bar{\boldsymbol{Z}}_0^\intercal\bar{\boldsymbol{Z}}_0\right)^{-1}\bar{\boldsymbol{Z}}_0^\intercal\boldsymbol{Z}_0, \boldsymbol{1}\left(\bar{\boldsymbol{Z}}_0^\intercal\bar{\boldsymbol{Z}}_0\right)^{-1}\bar{\boldsymbol{Z}}_0^\intercal\boldsymbol{Z}_0, \boldsymbol{1}\right)^\intercal\boldsymbol{Z}_0^\intercal\boldsymbol{u} + o_p(1),
\end{aligned}$$

and $\frac{1}{\sqrt{n}}\left(\boldsymbol{u}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{Z}, \boldsymbol{u}^\intercal\boldsymbol{M}\boldsymbol{Z}, \boldsymbol{u}^\intercal\bar{\boldsymbol{Z}}\right)^\intercal$ follows a multivariate normal distribution asymptotically. Since $\text{cov}\left(\boldsymbol{Z}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{u}, \bar{\boldsymbol{Z}}^\intercal\boldsymbol{u}\right) = \boldsymbol{0}$ and $\text{cov}\left(\boldsymbol{Z}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{u}, \boldsymbol{Z}^\intercal\boldsymbol{M}\boldsymbol{u}\right) = \boldsymbol{0}$, $\boldsymbol{Z}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{u}$ and $\left(\boldsymbol{Z}^\intercal\boldsymbol{M}\boldsymbol{u}, \bar{\boldsymbol{Z}}^\intercal\boldsymbol{u}\right)$ are asymptotically independent. This asymptotic independence can be used to show that $\boldsymbol{u}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{Z}\boldsymbol{Z}^\intercal\boldsymbol{M}\boldsymbol{u}$ and $\boldsymbol{u}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{Z}\boldsymbol{Z}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{u}$ are asymptotically uncorrelated, and $\boldsymbol{u}^\intercal(\boldsymbol{I} - \boldsymbol{M})\boldsymbol{Z}\boldsymbol{Z}^\intercal\boldsymbol{M}\boldsymbol{u}$ and $\boldsymbol{u}^\intercal\bar{\boldsymbol{Z}}\bar{\boldsymbol{Z}}^\intercal\boldsymbol{u}$ are asymptotically uncorrelated.

Thus,

$$\begin{aligned}\frac{1}{n}Q_\rho &= \frac{1}{n}(1-\rho)\boldsymbol{u}^\intercal(\boldsymbol{I}-\boldsymbol{M})\boldsymbol{Z}\boldsymbol{Z}^\intercal(\boldsymbol{I}-\boldsymbol{M})\boldsymbol{u} + \frac{1}{n}2(1-\rho)\boldsymbol{u}^\intercal(\boldsymbol{I}-\boldsymbol{M})\boldsymbol{Z}\boldsymbol{Z}^\intercal\boldsymbol{M}\boldsymbol{u} \\ &\quad + \frac{1}{n}\left[\rho p^2\bar{\boldsymbol{Z}}^\intercal\bar{\boldsymbol{Z}} + \frac{(1-\rho)\sum_{j=1}^{p}(\bar{\boldsymbol{Z}}^\intercal\boldsymbol{Z}_j)^2}{\bar{\boldsymbol{Z}}^\intercal\bar{\boldsymbol{Z}}}\right]\frac{\boldsymbol{u}^\intercal\bar{\boldsymbol{Z}}\bar{\boldsymbol{Z}}^\intercal\boldsymbol{u}}{\bar{\boldsymbol{Z}}^\intercal\bar{\boldsymbol{Z}}} + o_p(1) \\ &= (1-\rho)\sum\xi_k\chi^2_{1,k} + \tau(\rho)\eta_0 + o_p(1), \qquad (2\text{-}14)\end{aligned}$$

where $\tau(\rho) = \rho p^2\bar{\boldsymbol{Z}}_0^\intercal\bar{\boldsymbol{Z}}_0 + \frac{(1-\rho)\sum_{j=1}^{p}(\bar{\boldsymbol{Z}}_0^\intercal\boldsymbol{Z}_j^0)^2}{\bar{\boldsymbol{Z}}_0^\intercal\bar{\boldsymbol{Z}}_0}$, the $\chi^2_{1,k}$'s are independent and identically distributed χ^2_1 random variables, η_0 is a χ^2_1 random variable, and ξ_k is the k^{th} non-zero eigenvalue of $(\boldsymbol{I}-\boldsymbol{M}_0)\boldsymbol{Z}_0\boldsymbol{Z}_0^\intercal(\boldsymbol{I}-\boldsymbol{M}_0) + (I-\boldsymbol{M}_0)\boldsymbol{Z}_0\boldsymbol{Z}_0^\intercal\boldsymbol{M}_0 + \boldsymbol{M}_0\boldsymbol{Z}_0\boldsymbol{Z}_0^\intercal(I-\boldsymbol{M}_0)$. Hence Q_ρ can be decomposed into a weighted sum of two approximately independent random variables (Equation (2-14)), where the first random variable, $\sum\xi_k\chi^2_{1,k}$, is approximately independent of the second random variable, η_0, which is a χ^2_1 random variable.

Following Lee *and others* (2012), we obtain a p-value via one-dimensional integration:

Step 1: Set up a grid for ρ: $0 = \rho_1 < \rho_2 < \ldots < \rho_L = 1$. In the simulations and data application, we used $\rho_1 = 0, \rho_2 = 0.1, \rho_3 = 0.2, \cdots, \rho_{10} = 0.9, \rho_{11} = 1$.

Step 2: Compute $Q_{\rho_1}, \ldots, Q_{\rho_L}$, $\boldsymbol{Z}_0$, and $\boldsymbol{M}_0$, where $\boldsymbol{H}_0$ is replaced with $\boldsymbol{H}^{\widehat{\lambda}}$ in calculating $\boldsymbol{Z}_0$ and we plug in the ridge estimator of $\boldsymbol{\mu}$ at $\boldsymbol{\Delta}_0$.

Step 3: The mixture of chi-squares $\sum\xi_k\chi^2_{1,k}$ is approximated using $\sum_{k=1}^{m}\zeta_k\chi^2_{1,k}$ with an adjustment, where ζ_k is the k^{th} non-zero eigenvalue of $\boldsymbol{Z}_0^\intercal(I-\boldsymbol{M}_0)\boldsymbol{Z}_0$. Compute ζ_k for $k = 1, \cdots, m$ and $\tau(\rho_l)$ for $l = 1, \cdots, L$ and $\mu_Q = \sum_{k=1}^{m}\zeta_k$ and $\sigma_\psi = 2\sqrt{\text{trace}(\boldsymbol{Z}_0^\intercal\boldsymbol{M}_0\boldsymbol{Z}_0\boldsymbol{Z}_0^\intercal(\boldsymbol{I}-\boldsymbol{M}_0)\boldsymbol{Z}_0)}$ and $\sigma_Q = \sqrt{2\sum_{k=1}^{m}\zeta_k^2 + \sigma_\psi^2}$.

Step 4: Compute Q_{iSKAT}, p_{ρ_l} by characteristic function inversion (Davies, 1980) and $q_{\min}(\rho_l)$ using modified moment matching method (Liu *and others*, 2009), where $q_{\min}(\rho_l)$ is the $(1-Q_{\text{iSKAT}})^{\text{th}}$ quantile of the distribution of Q_{ρ_l}.

Step 5: Obtain a p-value for iSKAT via one-dimensional integration:

$$\text{p-value}_{\text{iSKAT}} = 1 - \int F(\delta(x)|\zeta)f(x|\chi^2_1)dx,$$

where $\delta(x) = [\min\{(q_{\min}(\rho_l) - \tau(\rho_l)x)/(1-\rho_l)\} - \mu_Q]\frac{\sqrt{\sigma_Q^2-\sigma_\psi^2}}{\sigma_Q} + \mu_Q$, $f(x|\chi^2_1)$ is the density function of a χ^2_1 random variable and $F(\delta(x)|\zeta)$ is the distribution function of a mixture of chi-squares $\sum_{k=1}^{m}\zeta_k\chi^2_k$ obtained by characteristic function inversion method (Davies, 1980).

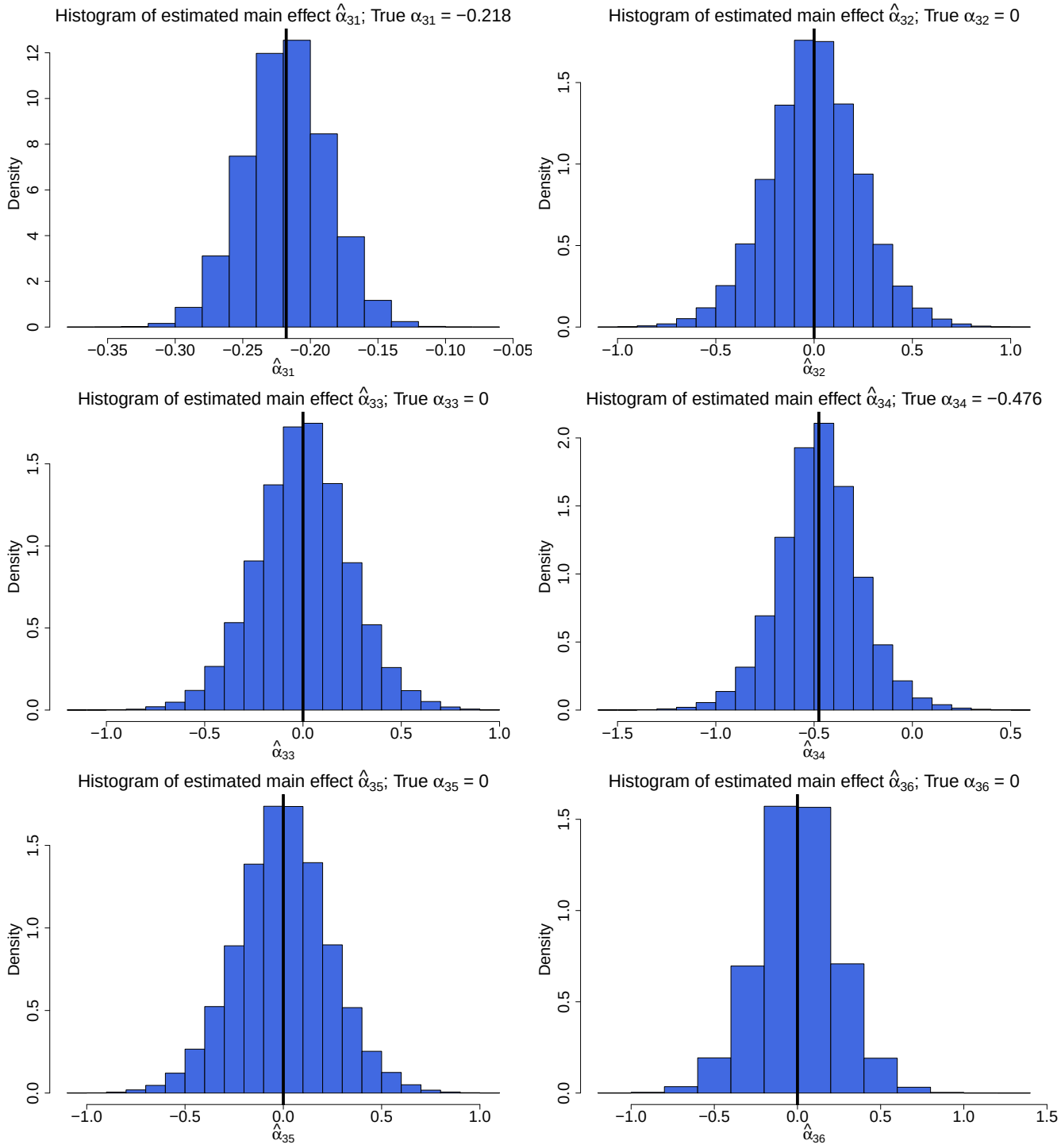
Web Appendix 3. Additional Simulation Studies Based on the CoLaus Dataset

3.1. Estimated Rare Variants Main Effects from iSKAT

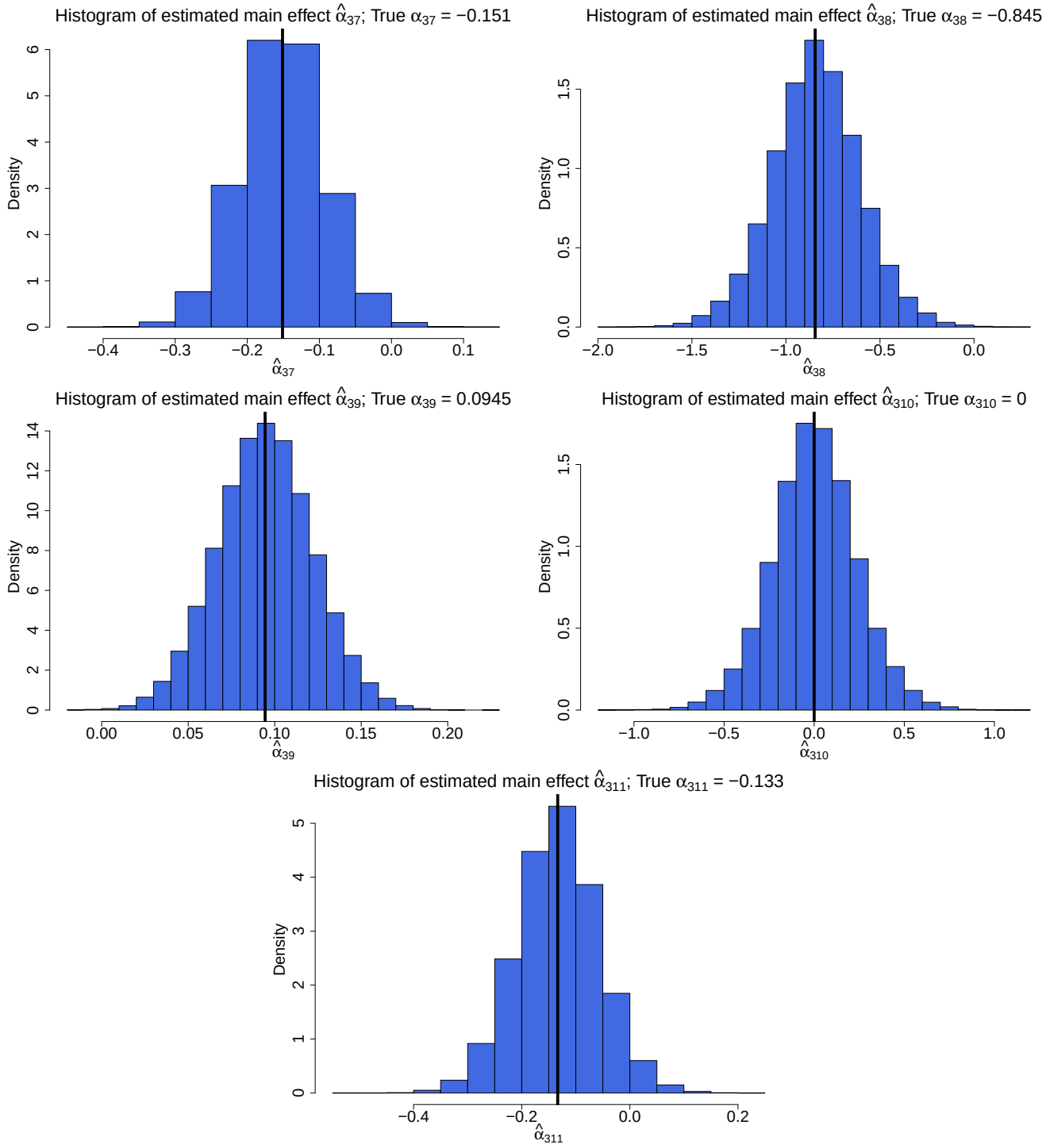
In the top two panels of main manuscript Table 1, we report empirical Type 1 error rates for a continuous outcome when there are rare variants main effects, for $n = 1945$ and $n = 4000$ respectively. The (true) rare variants main effect sizes are $\boldsymbol{\alpha}_3 = (-0.218, 0, 0, -0.476, 0, 0, -0.151, -0.845, 0.0945, 0, -0.133)^\intercal$ in the simulations (main manuscript Section 5). iSKAT obtains estimates of rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ using weighted ridge regression (Web Appendix 2.1.). For the simulations described in Section 5 of the main manuscript and reported in the top two panels of main manuscript Table 1, the histograms of estimated rare variants main effects for each of the rare variants are given in Web Figures 1-2 and 3-4, for $n = 1945$ and $n = 4000$ respectively. The average of these estimated main effects are also reported in the third and fourth columns of Web Table 1, for $n = 1945$ and $n = 4000$ respectively. The estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ are similar to the true rare variants main effects $\boldsymbol{\alpha}_3$.

Web Table 1: Average estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ from iSKAT, for continuous outcome when there are rare variants main effects for $n = 1945$ (third column) and $n = 4000$ (fourth column) respectively. True rare variants main effect sizes $\boldsymbol{\alpha}_3$ are given in the second column. Datasets were generated by sampling the genotypes and covariates jointly with replacement from the CoLaus dataset to preserve the association between $\boldsymbol{G}$ and E. Note that the p-value for the association between $\boldsymbol{G}$ and E in the CoLaus dataset was 0.042, which suggests plausible $\boldsymbol{G} - E$ dependence. Estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ are similar to the true rare variants main effects $\boldsymbol{\alpha}_3$.

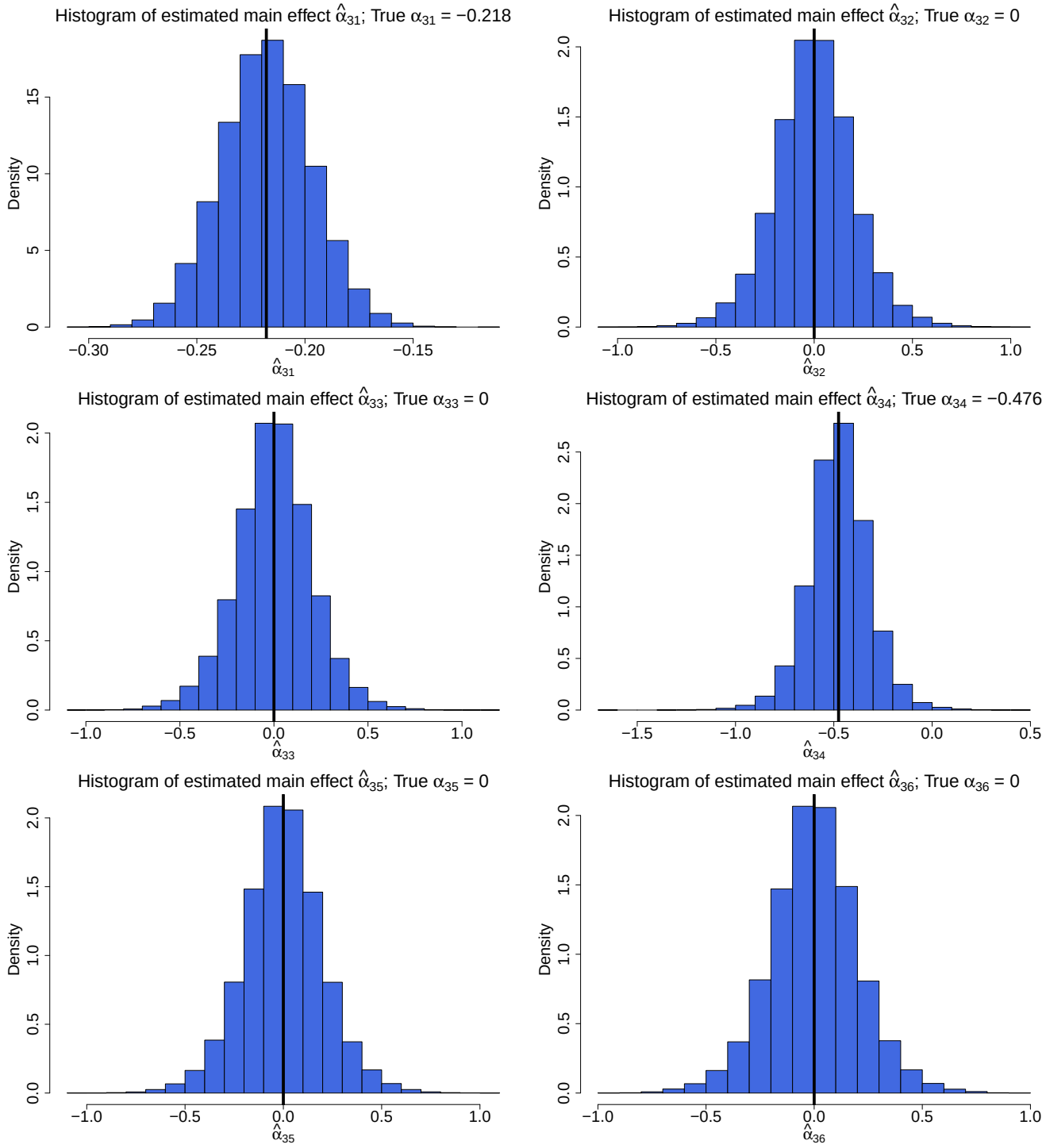
Variant	True α_{3j}	Average $\hat{\alpha}_{3j}$ when $n = 1945$	Average $\hat{\alpha}_{3j}$ when $n = 4000$
chr3:188053510	-2.18e-01	-2.18e-01	-2.18e-01
chr3:188053533	0	1.86e-04	-2.39e-04
chr3:188053663	0	-7.14e-04	-1.95e-04
chr3:188053705	-4.76e-01	-4.73e-01	-4.74e-01
chr3:188053732	0	9.22e-05	-3.81e-04
chr3:188054673	0	-3.02e-04	-3.95e-04
chr3:188054720	-1.51e-01	-1.51e-01	-1.51e-01
chr3:188054724	-8.45e-01	-8.39e-01	-8.38e-01
chr3:188054783	9.45e-02	9.45e-02	9.44e-02
chr3:188055047	0	1.17e-03	-2.79e-04
chr3:188055252	-1.33e-01	-1.33e-01	-1.33e-01



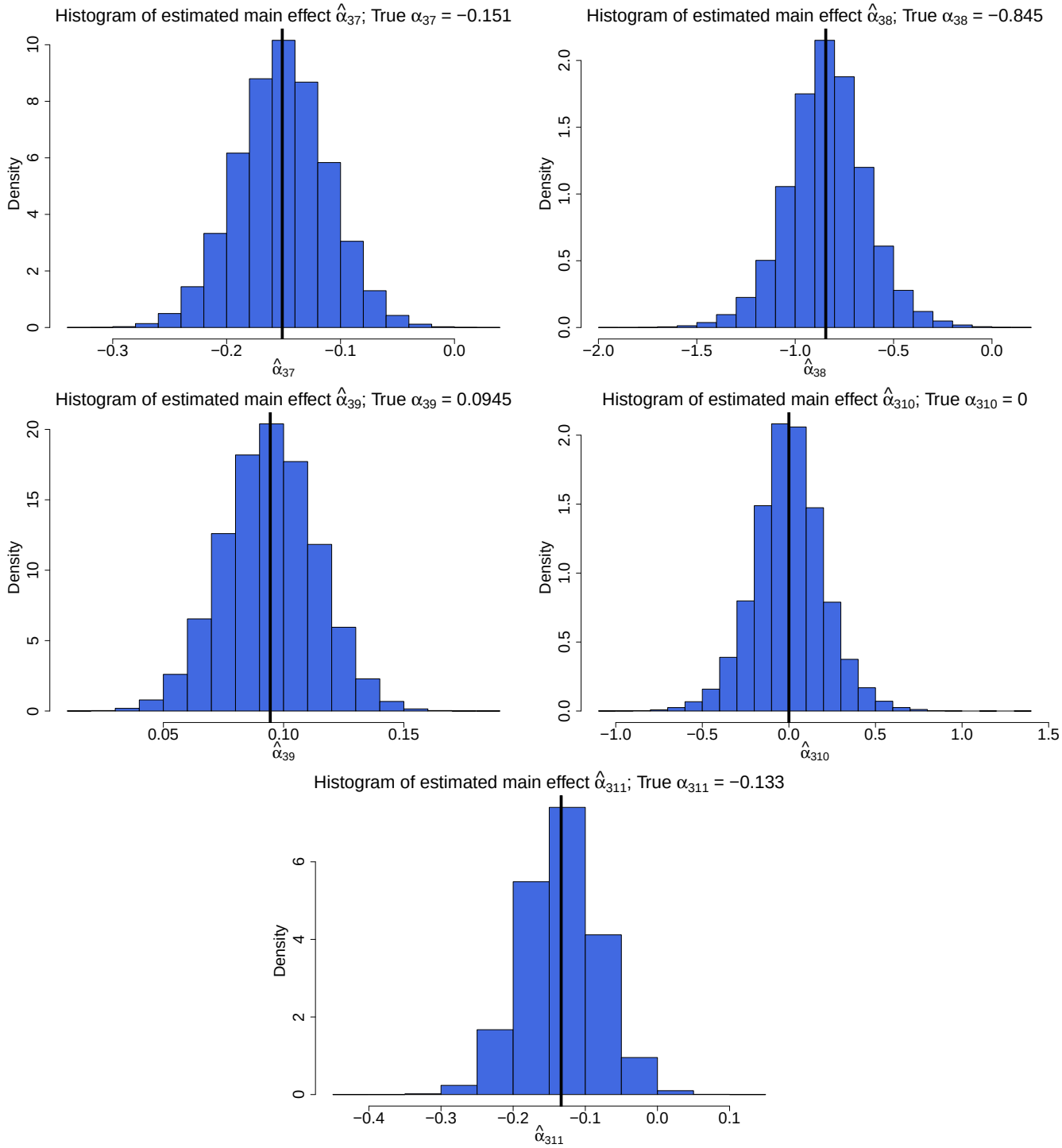
Web Figure 1: Histograms of estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ from iSKAT, for continuous outcome when there are rare variants main effects, for $n = 1945$. True rare variants main effect sizes $\boldsymbol{\alpha}_3$ are indicated with vertical lines. This figure gives $\hat{\alpha}_{3j}$ for $j = 1 - 6$ while Web Figure 2 gives $\hat{\alpha}_{3j}$ for $j = 7 - 11$. Datasets were generated by sampling the genotypes and covariates jointly with replacement from the CoLaus dataset to preserve the association between $\boldsymbol{G}$ and E. Note that the p-value for the association between $\boldsymbol{G}$ and E in the CoLaus dataset was 0.042, which suggests plausible $\boldsymbol{G} - E$ dependence. Estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ are similar to the true rare variants main effects $\boldsymbol{\alpha}_3$.



Web Figure 2: Histograms of estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ from iSKAT, for continuous outcome when there are rare variants main effects, for $n = 1945$. True rare variants main effect sizes $\boldsymbol{\alpha}_3$ are indicated with vertical lines. This figure gives $\hat{\alpha}_{3j}$ for $j = 7 - 11$ while Web Figure 1 gives $\hat{\alpha}_{3j}$ for $j = 1 - 6$. Datasets were generated by sampling the genotypes and covariates jointly with replacement from the CoLaus dataset to preserve the association between $\boldsymbol{G}$ and E. Note that the p-value for the association between $\boldsymbol{G}$ and E in the CoLaus dataset was 0.042, which suggests plausible $\boldsymbol{G} - E$ dependence. Estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ are similar to the true rare variants main effects $\boldsymbol{\alpha}_3$.



Web Figure 3: Histograms of estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ from iSKAT, for continuous outcome when there are rare variants main effects, for $n = 4000$. True rare variants main effect sizes $\boldsymbol{\alpha}_3$ are indicated with vertical lines. This figure gives $\hat{\alpha}_{3j}$ for $j = 1 - 6$ while Web Figure 4 gives $\hat{\alpha}_{3j}$ for $j = 7 - 11$. Datasets were generated by sampling the genotypes and covariates jointly with replacement from the CoLaus dataset to preserve the association between $\boldsymbol{G}$ and E. Note that the p-value for the association between $\boldsymbol{G}$ and E in the CoLaus dataset was 0.042, which suggests plausible $\boldsymbol{G} - E$ dependence. Estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ are similar to the true rare variants main effects $\boldsymbol{\alpha}_3$.



Web Figure 4: Histograms of estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ from iSKAT, for continuous outcome when there are rare variants main effects for, $n = 4000$. True rare variants main effect sizes $\boldsymbol{\alpha}_3$ are indicated with vertical lines. This figure gives $\hat{\alpha}_{3j}$ for $j = 7 - 11$ while Web Figure 3 gives $\hat{\alpha}_{3j}$ for $j = 1 - 6$. Datasets were generated by sampling the genotypes and covariates jointly with replacement from the CoLaus dataset to preserve the association between $\boldsymbol{G}$ and E. Note that the p-value for the association between $\boldsymbol{G}$ and E in the CoLaus dataset was 0.042, which suggests plausible $\boldsymbol{G} - E$ dependence. Estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ are similar to the true rare variants main effects $\boldsymbol{\alpha}_3$.

3.2. Empirical p-values for iSKAT

In Web Appendix 2.3., we described how a p-value for iSKAT is obtained using the asympotic decomposition in Equation (2-14) and one-dimensional integration. To evaluate the performance of the asymptotic p-values (Web Appendix 2.3.), in this section, we describe how an empirical p-value for iSKAT can be obtained, before comparing the empirical and asymptotic p-values in Web Appendix 3.3.

Step 1: Set up a grid for ρ: $0 = \rho_1 < \rho_2 < \ldots < \rho_L = 1$. We used $\rho_1 = 0, \rho_2 = 0.1, \rho_3 = 0.2, \cdots, \rho_{10} = 0.9, \rho_{11} = 1$.

Step 2: Compute $Q_{\rho_1}, \ldots, Q_{\rho_L}$ using Equation (6) in the main manuscript.

Step 3: Compute p_{ρ_l} by characteristic function inversion (Davies, 1980) as described in Web Appendix 2.2. Compute Q_{iSKAT} (Equation (8) in the main manuscript).

Step 4: To obtain the empirical null distribution of the test statistic Q_{iSKAT}, using the notation in Section 2 of the main manuscript, let the observed data be $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{E}, \boldsymbol{G})$.

- Fit the null model (Equation (7) of the main manuscript) to the observed data $(\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{E}, \boldsymbol{G})$ using weighted ridge regression (Web Appendix 2.1.) and obtain $\boldsymbol{\mu}(\widehat{\boldsymbol{\alpha}})$. For a continuous outcome, we also obtain $\widehat{\sigma}^2$, an estimate of the variance from the residuals.
- For each re-sample ($b = 1, \cdots, B$):
  1. For continuous outcome (linear regression), generate $\boldsymbol{Y}^b_{\text{new}} \sim \text{Normal}\left(\boldsymbol{\mu}(\widehat{\boldsymbol{\alpha}}), \widehat{\sigma}^2\right)$. For binary outcome (logistic regression), generate $\boldsymbol{Y}^b_{\text{new}} \sim \text{Binomial}\left(1, \boldsymbol{\mu}(\widehat{\boldsymbol{\alpha}})\right)$.
  2. Compute Q^b_{iSKAT} from $\left(\boldsymbol{Y}^b_{\text{new}}, \boldsymbol{X}, \boldsymbol{E}, \boldsymbol{G}\right)$ using Steps 1-3 above.

Step 5: Obtain an empirical p-value for iSKAT by comparing the observed test statistic Q_{iSKAT} (from Step 3) with the empirical null distribution (from Step 4):

$$\text{p-value}_{\text{iSKAT,empirical}} = \frac{1}{B}\sum_{b=1}^{B} I(Q^b_{\text{iSKAT}} \leq Q_{\text{iSKAT}}).$$

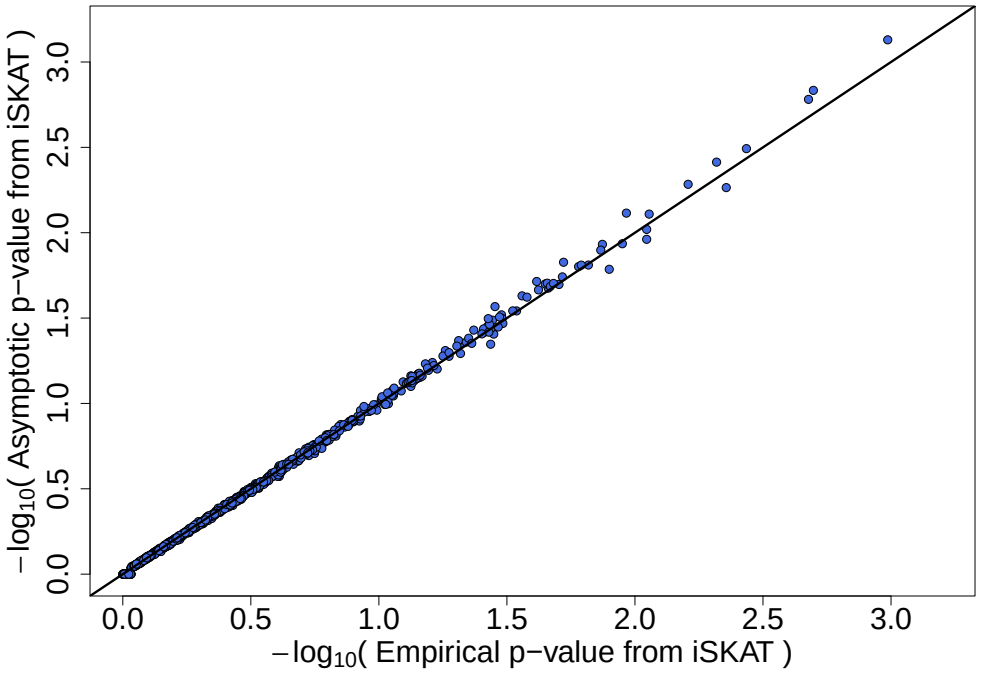
3.3. Simulations to compare empirical and asymptotic p-values for iSKAT

To evaluate the performance of the asymptotic p-values (obtained using one-dimensional integration as described in Web Appendix 2.3.), we compared the asymptotic p-values (Web Appendix 2.3.) with the empirical p-values (obtained using the description in Web Appendix 3.2.). We consider simulations similar to those presented in Section 5 of the main text. Empirical p-values are obtained using $B = 5000$ re-samples. If the empirical p-value was less than 2.5×10^{-3}, we increased the number of re-samples to $B = 1 \times 10^5$. Datasets were generated by sampling the genotypes and covariates (including the environmental factor) jointly with replacement from the CoLaus dataset. We consider a sample size of $n = 1945$:

$$Y_i = \boldsymbol{X}_i^\intercal \boldsymbol{\alpha}_1 + E_i \alpha_2 + \boldsymbol{G}_i^\intercal \boldsymbol{\alpha}_3 + E_i \boldsymbol{G}_i^\intercal \boldsymbol{\beta} + \epsilon_i,$$

where $\boldsymbol{\alpha}_1 = (3.6, -0.030, -1.4, 8.3, -4.1, 2.2, 0.005, -0.015, -0.0056, 0.0069, \ -0.033, 0.15)^\intercal$, $\alpha_2 = 0.015$, $\boldsymbol{\alpha}_3 = (-0.218, 0, 0, -0.476, 0, 0, -0.151, -0.845, 0.0945, 0, -0.133)^\intercal$, $\boldsymbol{\beta} = \mathbf{0}$ and $\epsilon_i \sim N(0, 0.27)$.

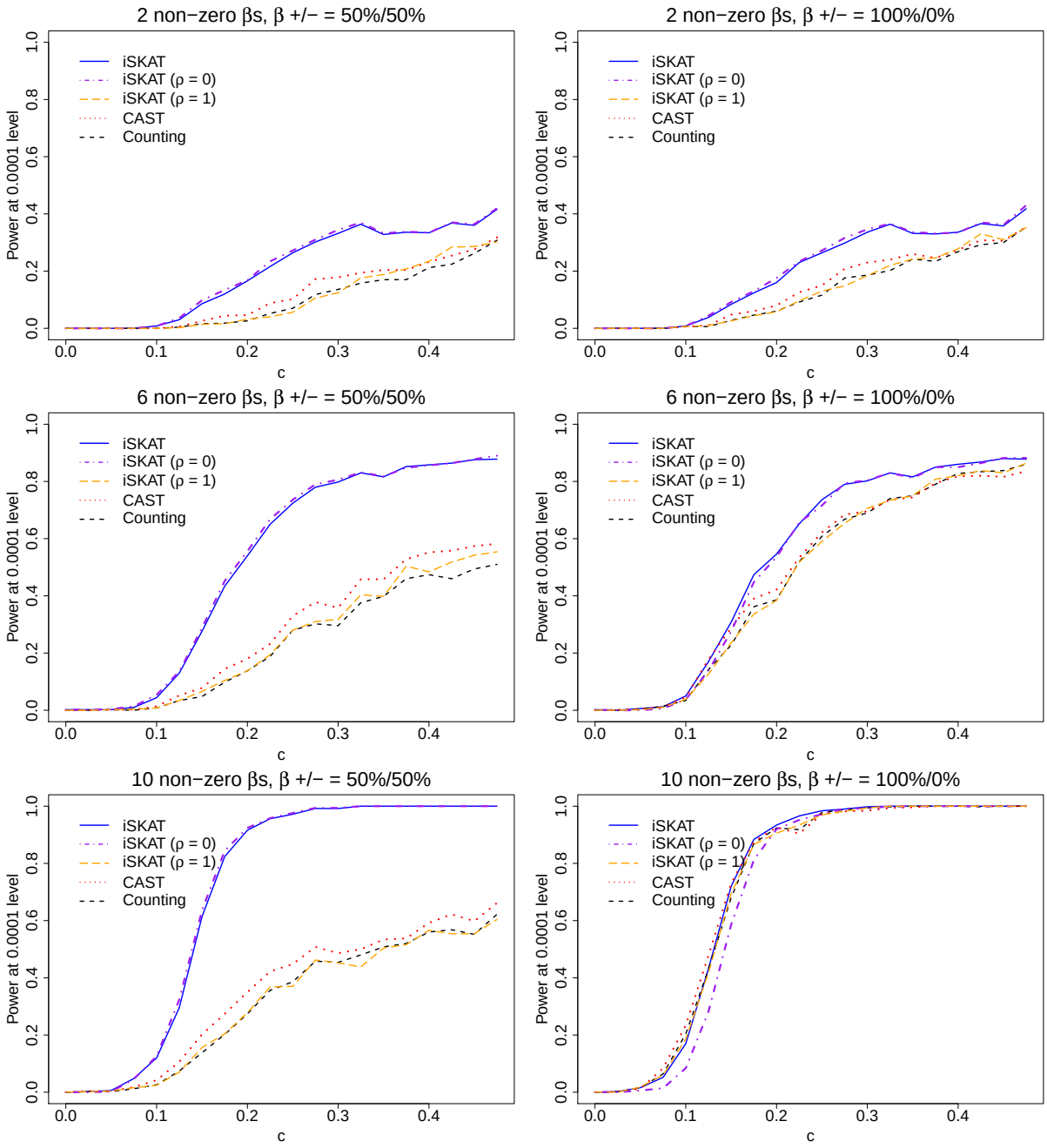
We compared the asymptotic and empirical p-values for 1000 simulations. The plot of asymptotic p-values against the empirical p-values are given in Web Figure 5. Both methods give very similar p-values. However, obtaining p-values empirically is computationally very intensive, especially for small p-values.



Web Figure 5: Plot of $-\log_{10}$(Asymptotic p-value) (vertical axis) against $-\log_{10}$(Empirical p-value) (horizontal axis) from iSKAT, for continuous outcome and there are main effects for $n = 1945$. Datasets were generated by sampling the genotypes and covariates jointly with replacement from the CoLaus dataset to preserve the association between $\boldsymbol{G}$ and E. Note that the p-value for the association between $\boldsymbol{G}$ and E in the CoLaus dataset was 0.042, which suggests plausible $\boldsymbol{G} - E$ dependence. Empirical (horizontal axis) and asymptotic p-values (vertical axis) are similar.

3.4. Additional power simulations based on the CoLaus Dataset

Web Figure 6 gives the empirical power curves at $n = 4000$ for the CoLaus Dataset. The simulation details are given in Section 5 of the main text. Analogous power curves at $n = 2000$ are presented in Figure 1 of the main manuscript.



Web Figure 6: Empirical power curves for $n = 4000$ at $\alpha = 0.0001$ level of significance for testing rare variant GE interaction effects on a continuous outcome when there are no main effects - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 2 non-zero β_j's; Middle panel - 6 non-zero β_j's; Bottom panel - 10 non-zero β_j's. Left panel - 50% of β_j's are positive; Right panel -100% of β_j's are positive. In each plot, we set the magnitudes of the non-zero β_j's as $|\beta_j| = c$, and increased c from zero until 0.475. Datasets were generated by sampling the genotypes and covariates jointly with replacement from the CoLaus dataset to preserve the association between $\boldsymbol{G}$ and E. Note that the p-value for the association between $\boldsymbol{G}$ and E in the CoLaus dataset was 0.042, which suggests plausible $\boldsymbol{G} - E$ dependence.

Web Appendix 4. Simulation Studies Based on Coalescent Model

In the main text and Web Appendix 3., we present simulation studies based on the CoLaus dataset. In this section, we provide additional simulation results from more general settings. We evaluate the performance of (1) iSKAT, (2) iSKAT with $\rho = 0$, (3) iSKAT with $\rho = 1$, (4) CAST and (5) Counting. These five methods are also described in the main text.

Sequence genotypes are generated from a coalescent model (Schaffner *and others*, 2005) by mimicking the linkage disequilibrium pattern for European ancestry. 10,000 haplotypes over a 10^6 base pairs region are generated and to form genotypes, 2 haplotypes are randomly selected. Since the average length of the exome region of a gene is around 3000 base pairs, we randomly select regions with 3000 base pairs in length, and test for rare variants by environment interactions in all simulation settings. We consider both continuous and binary outcomes and two different sample sizes $n = 2000, 4000$. For binary outcome, case-control sampling is used to obtain 50% cases and 50% controls. To generate the outcomes, we use:

$$g(\mu_i) = \alpha_0 - 0.2X_{1i} - 0.2X_{2i} + 0.4E_i + \boldsymbol{G}_i^{\intercal}\boldsymbol{\alpha}_3 + E_i\boldsymbol{G}_i^{\intercal}\boldsymbol{\beta}, \tag{4-15}$$

where $g(\cdot)$ is the identity and logit functions for continuous and binary outcomes respectively, and $X_{1i} \sim N(1,1)$, $X_{2i} \sim N(1,1)$. For continuous outcome, we assume $\alpha_0 = 0$, $\mathrm{Var}(Y_i|\boldsymbol{X}_i, E_i, \boldsymbol{G}_i) = 1$ and Y_i is normally distributed. For binary outcome, we assume $\alpha_0 = \log\left(\frac{0.01}{0.99}\right)$. We generate E_i from:

$$\mathrm{logit}\left[P(E_i = 1|\boldsymbol{G}_i)\right] = \boldsymbol{G}_i^{\intercal}\boldsymbol{\gamma}_3. \tag{4-16}$$

We consider two distinct scenarios when (i) $\boldsymbol{G}$ and E are independent, i.e. $\boldsymbol{\gamma}_3 = \mathbf{0}$ in Equation (4-16) and (ii) $\boldsymbol{G}$ and E are dependent, i.e. $\boldsymbol{\gamma}_3 \neq \mathbf{0}$ in Equation (4-16).

4.1. Size Simulations

To evaluate the empirical Type 1 error rates, we set $\boldsymbol{\beta} = \mathbf{0}$ in Equation (4-15). We consider the two cases when there are (a) no main effects of the rare variants, i.e. $\boldsymbol{\alpha}_3 = \mathbf{0}$ in Equation (4-15) and (b) main effects of the rare variants, i.e. $\boldsymbol{\alpha}_3 \neq \mathbf{0}$ in Equation (4-15).

When there are (a) no main effects of the rare variants and (i) $\boldsymbol{G}$ and E are independent, we set $\boldsymbol{\alpha}_3 = \boldsymbol{\gamma}_3 = \mathbf{0}$. When there are (a) no main effects of the rare variants and (ii) $\boldsymbol{G}$ and E are dependent, we set $\boldsymbol{\alpha}_3 = \mathbf{0}$ and consider the cases when 20%, 50% or 80% of the γ_{3j}'s are non-zero, and set the non-zero γ_{3j}'s as $|\gamma_{3j}| = c|\log_{10} \text{MAF}_j|$. We consider the scenarios when all γ_{3j}'s are positive and when 50% of the γ_{3j}'s are positive, while varying c. When there are (b) main effects of the rare variants, we consider the cases when 20%, 50% or 80% of the α_{3j}'s are non-zero, and set the non-zero α_{3j}'s as $|\alpha_{3j}| = c|\log_{10} \text{MAF}_j|$ and look at the scenarios when all α_{3j}'s are positive and when 50% of the α_{3j}'s are positive, while varying c. When there are (b) main effects of the rare variants and (i) $\boldsymbol{G}$ and E are independent, $\boldsymbol{\alpha}_3$ is generated as described and we set $\boldsymbol{\gamma}_3 = \mathbf{0}$. When there are (b) main effects of the rare variants and (ii) $\boldsymbol{G}$ and E are dependent, $\boldsymbol{\alpha}_3$ is generated as described and we set $\boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$.

The empirical Type 1 error rates when there are (a) no main effects of the rare variants, are given in Web Table 2 ($\boldsymbol{G} \perp E$, $n = 2000, 4000$) and Web Figures 7 ($\boldsymbol{G} \not\perp E$, $n = 2000$), 8 ($\boldsymbol{G} \not\perp E$, $n = 4000$) for continuous outcome; and Web Table 3 ($\boldsymbol{G} \perp E$, $n = 2000, 4000$) and Web Figures 17 ($\boldsymbol{G} \not\perp E$, $n = 2000$), 18 ($\boldsymbol{G} \not\perp E$, $n = 4000$) for binary outcome. Since there are no main effects of the rare variants, all five methods have correct Type 1 error rates.

The empirical Type 1 error rates when there are (b) main effects of the rare variants, are given in Web Figures 9 ($\boldsymbol{G} \perp E$, $n = 2000$), 10 ($\boldsymbol{G} \perp E$, $n = 4000$), 11 ($\boldsymbol{G} \not\perp E$, $n = 2000$) and 12 ($\boldsymbol{G} \not\perp E$, $n = 4000$) for continuous outcome; and in Web Figures 19 ($\boldsymbol{G} \perp E$, $n = 2000$), 20 ($\boldsymbol{G} \perp E$, $n = 4000$), 21 ($\boldsymbol{G} \not\perp E$, $n = 2000$), 22 ($\boldsymbol{G} \not\perp E$, $n = 4000$) for binary outcome. In all of these figures, with the exception of Web Figures 19 ($\boldsymbol{G} \perp E$, $n = 2000$) and 20 ($\boldsymbol{G} \perp E$, $n = 4000$), the burden tests can have inflated Type 1 error rates since there are main effects of the rare variants. In Web Figures 19 ($\boldsymbol{G} \perp E$, $n = 2000$) and 20 ($\boldsymbol{G} \perp E$, $n = 4000$), the burden tests also have correct Type 1 error rates since as discussed in the main text, this is a special case when $\boldsymbol{G} \perp E$ and we have a binary outcome from a logistic model with rare disease assumption.

Web Figures 11-12, 21-22 depict the scenario when there are (b) main effects of the rare variants and (ii) $\boldsymbol{G}$ and E are dependent. In these four figures, we have $\boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$ and the magnitudes of the non-zero α_{3j}'s and γ_{3j}'s are such that $|\alpha_{3j}| = |\gamma_{3j}| = c|\log_{10} \text{MAF}_j|$. As c increases in these figures, *both* $\boldsymbol{G} - Y$ and $\boldsymbol{G} - E$ associations increase, due to the increasing magnitudes of α_{3j}'s and γ_{3j}'s respectively. The bias of the burden tests increase as c increases in these four figures, due to *both* increasing $\boldsymbol{G} - Y$ and $\boldsymbol{G} - E$ associations. To further study how (1) the main effects of the rare variants on the outcome ($\boldsymbol{G} - Y$ association) and (2) $\boldsymbol{G} - E$ dependence, *individually* affect the bias of the burden tests, we ran additional simulations when there are (b) main effects of the rare variants and (ii) $\boldsymbol{G} - E$ dependent, but holding (1) magnitudes of the non-zero γ_{3j}'s fixed or (2) magnitudes of the non-zero α_{3j}'s

fixed.

To study how the (1) main effects of the rare variants on the outcome ($\boldsymbol{G} - Y$ association) affect the bias of the burden tests, we consider when 20%, 50% or 80% of the α_{3j}'s are non-zero, and set the non-zero α_{3j}'s as $|\alpha_{3j}| = c|\log_{10} \text{MAF}_j|$ and look at the scenarios when all α_{3j}'s are positive and when 50% of the α_{3j}'s are positive, while varying c. However, now we set $\gamma_{3j} = 0.4\text{sgn}(\alpha_{3j})|\log_{10} \text{MAF}_j|$ for $j = 1, \cdots, p$, where $\text{sgn}(\cdot)$ is the sign function, i.e. the magnitudes of the non-zero γ_{3j}'s do not depend on c. The empirical Type 1 error rates are given in Web Figures 13-14 for continuous outcome and Web Figures 23-24 for binary outcome respectively. In Web Figures 13-14, 23-24, the bias of the burden tests increase with increasing c (increasing $\boldsymbol{G} - Y$ association) when $\boldsymbol{G}$ and E are dependent, which is what we expect from the calculation in Web Appendix 1.3.

To investigate how the bias of the burden tests is affected by (2) $\boldsymbol{G} - E$ dependence, we consider when 20%, 50% or 80% of the γ_{3j}'s are non-zero, and set the non-zero γ_{3j}'s as $|\gamma_{3j}| = c|\log_{10} \text{MAF}_j|$ and look at the scenarios when all γ_{3j}'s are positive and when 50% of the γ_{3j}'s are positive, while varying c. We set $\alpha_{3j} = 0.4\text{sgn}(\gamma_{3j})|\log_{10} \text{MAF}_j|$ for $j = 1, \cdots, p$, where $\text{sgn}(\cdot)$ is the sign function, i.e. the magnitudes of the non-zero α_{3j}'s do not depend on c. The empirical Type 1 error rates are given in Web Figures 15-16 for continuous outcome and Web Figures 25-26 for binary outcome respectively. In Web Figures 15-16, 25-26, the bias of the burden tests increase with increasing c (increasing $\boldsymbol{G} - E$ dependence) when $\boldsymbol{G}$ and Y are associated, which is what we expect from the calculation in Web Appendix 1.3.

4.2. Power Simulations

To study the empirical power, we consider the case when there are (a) no main effects of the rare variants i.e. $\boldsymbol{\alpha}_3 = \boldsymbol{0}$ in Equation (4-15) and (i) $\boldsymbol{G}$ and E are independent i.e. $\boldsymbol{\gamma}_3 = \boldsymbol{0}$ in Equation (4-16). We consider the scenarios when 20%, 50% or 80% of the β_j's in Equation (4-15) are non-zero, and set the non-zero β_j's as $|\beta_j| = c|\log_{10} \text{MAF}_j|$ and look at the scenarios when all β_j's are positive and when 50% of the β_j's are positive, while varying c. The power curves are given in Web Figures 27 ($n = 2000$) and 28 ($n = 4000$) for continuous outcome; and Web Figures 29 ($n = 2000$) and 30 ($n = 4000$) for binary outcome. iSKAT generally has power at least comparable to or outperforms the burden tests. In all the plots except when all β_j's are positive and the majority of the β_j's are non-zero, iSKAT has power comparable to iSKAT with $\rho = 0$. In the plots when all β_j's are positive and the majority of the β_j's are non-zero, iSKAT has power comparable to iSKAT with $\rho = 1$.

4.3. Simulations to compare estimation of null model with and without penalization

In Web Appendix 2.1., we describe how weighted ridge regression is used to estimate the null model (Equation (7) of the main manuscript), and noted that $\widehat{\boldsymbol{\alpha}}^{\lambda=0}$ corresponds to the usual maximum likelihood estimator, where λ is the ridge parameter. In this section we compare Type 1 error rates and power of iSKAT when (1) the null model is estimated without penalty ($\lambda = 0$) and (2) the null model is estimated with penalty (λ selected using generalized cross-validation as described in Web Appendix 2.1.).

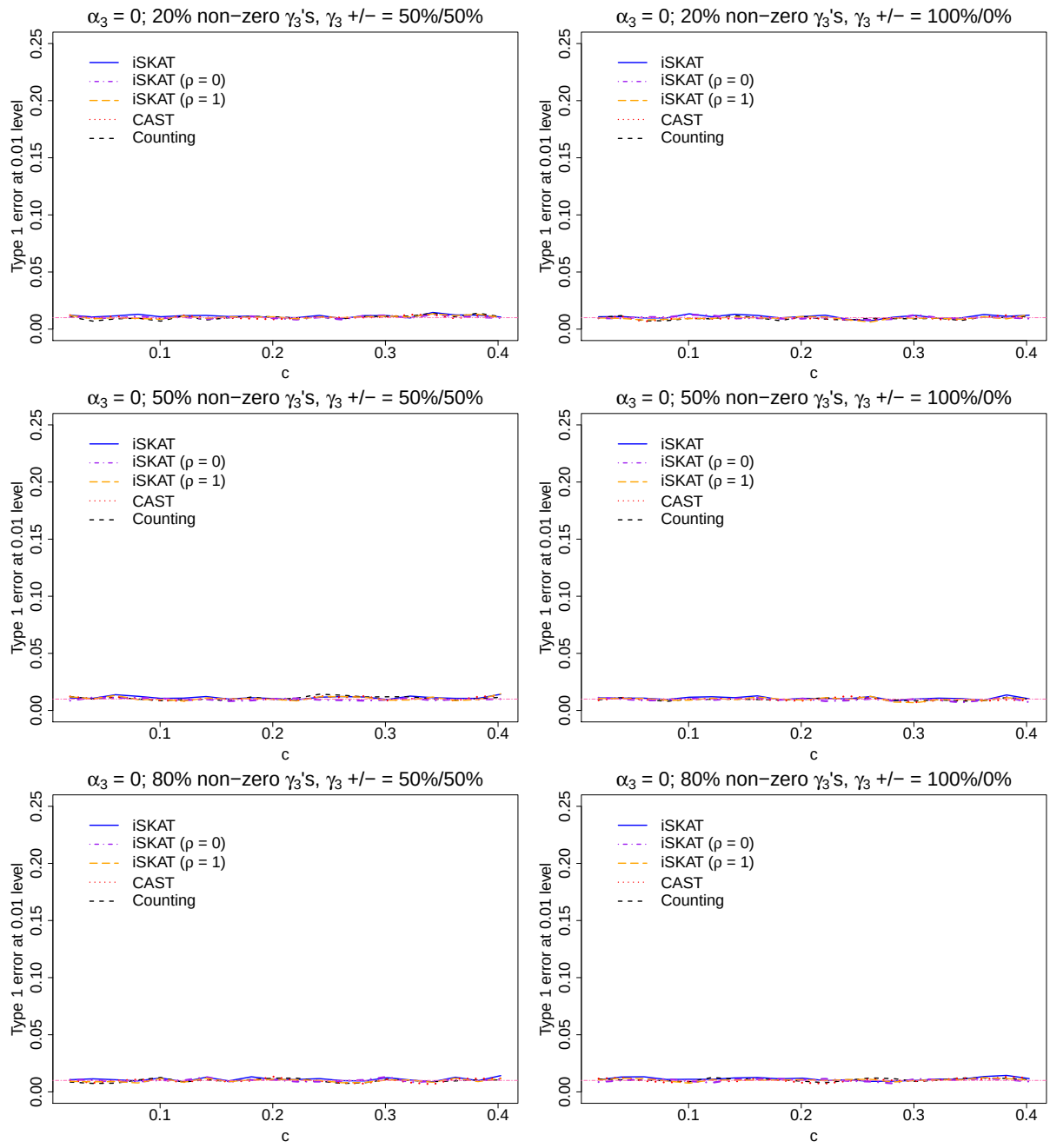
We consider the scenario similar to that in Web Figure 11, when we have a continuous outcome and there are (b) main effects of the rare variants and (ii) $\boldsymbol{G}$ and E are dependent. Similar to before, we consider the cases when 20%, 50% or 80% of the α_{3j}'s are non-zero, and set the non-zero α_{3j}'s as $|\alpha_{3j}| = c|\log_{10} \text{MAF}_j|$ and look at the scenarios when all α_{3j}'s are positive and when 50% of the α_{3j}'s are positive, while varying c. We set $\boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$. We evaluated 20 values of c from 0.02 to 0.4 in equal intervals, and had 5000 simulations for each value of c, giving a total of 10^5 simulations. We report the empirical Type 1 error rates and power averaged over 10^5 simulations. To evaluate Type 1 error and power, we set $\boldsymbol{\beta} = \mathbf{0}$ and $\boldsymbol{\beta} = \boldsymbol{\alpha}_3$ respectively.

iSKAT with the null model estimated without penalization ($\lambda = 0$) did not converge for 71% out of the 10^5 simulations due to singularity or other numerical problems. We report the empirical Type 1 error rates and power for the remainder 29% of simulations that converged in Web Tables 4 and 5 respectively. The Type 1 error rates and power of both approaches in the remaining 29% of simulations are similar.

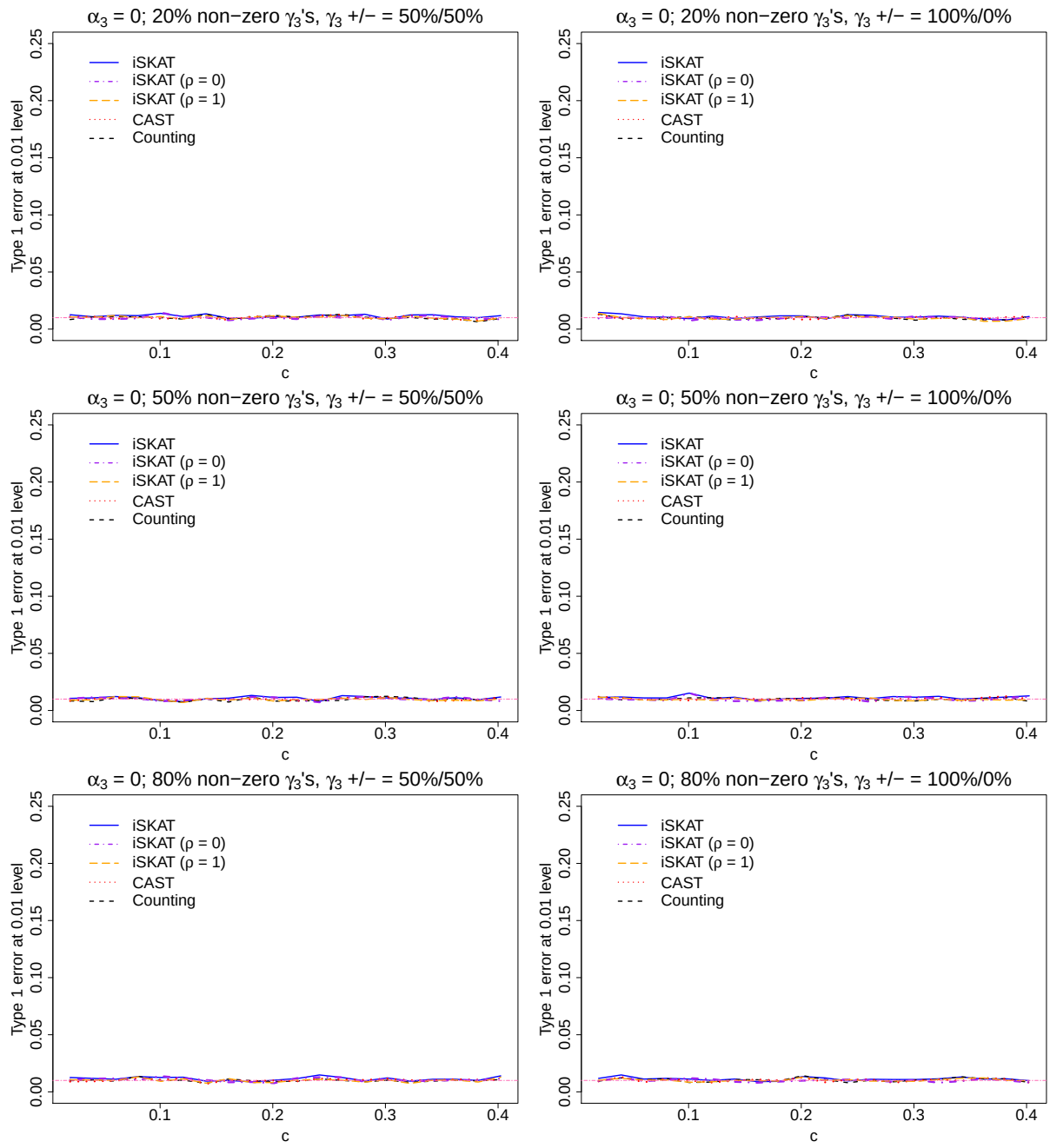
Web Table 2: Empirical Type 1 error rates for **continuous** outcome when $\boldsymbol{G}$ and E are **independent** and there are **no main** effects (i.e. $\boldsymbol{\alpha}_3 = \boldsymbol{\gamma}_3 = \mathbf{0}$) for $n = 2000$ (top panel) and $n = 4000$ (bottom panel) respectively. Since there are no main effects for rare variants, all five methods have correct Type 1 error rates.

	α-level	iSKAT	iSKAT ($\rho = 0$)	iSKAT ($\rho = 1$)	CAST	Counting
$n = 2000$	1e-02	1.16e-02	1.03e-02	1.06e-02	1.07e-02	1.03e-02
	1e-03	1.16e-03	9.00e-04	1.14e-03	1.16e-03	1.20e-03
	1e-04	1.30e-04	8.00e-05	1.30e-04	9.00e-05	2.00e-04

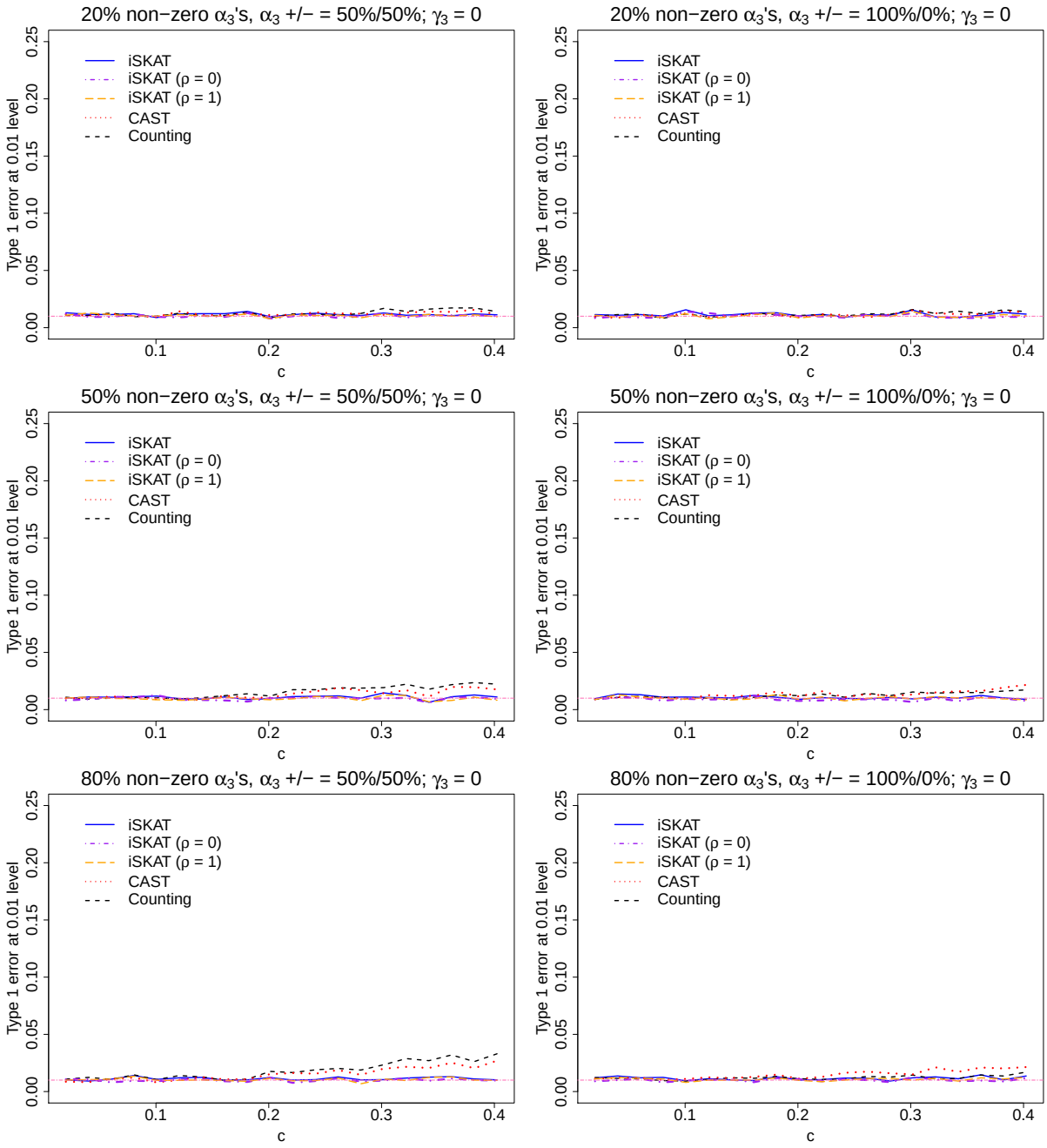
	α-level	iSKAT	iSKAT ($\rho = 0$)	iSKAT ($\rho = 1$)	CAST	Counting
$n = 4000$	1e-02	1.11e-02	1.03e-02	9.89e-03	1.00e-02	1.00e-02
	1e-03	1.49e-03	1.12e-03	1.14e-03	1.13e-03	1.19e-03
	1e-04	1.20e-04	1.10e-04	1.00e-04	8.00e-05	1.40e-04



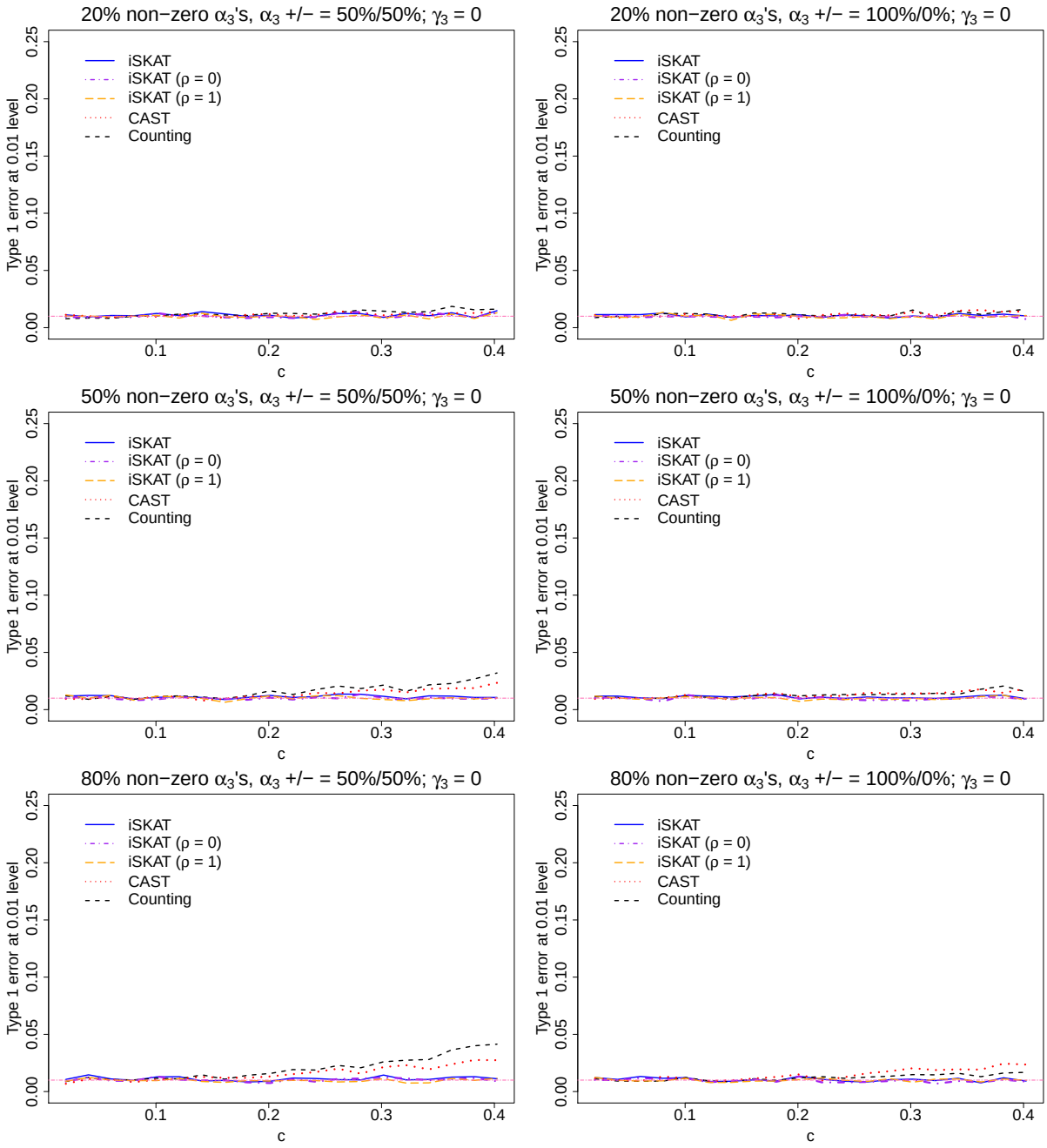
Web Figure 7: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **no main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero γ_{3j}'s; Middle panel - 50% non-zero γ_{3j}'s; Bottom panel - 80% non-zero γ_{3j}'s. Left panel - 50% of γ_{3j}'s are positive; Right panel - 100% of γ_{3j}'s are positive. $\boldsymbol{\alpha_3 = 0}$. $\boldsymbol{\gamma_3}$ controls the association between $\boldsymbol{G}$ and E. **As** c **increases, the** $\boldsymbol{G} - E$ **association increases.** Since there are no main effects for rare variants, all five methods have correct Type 1 error rates.



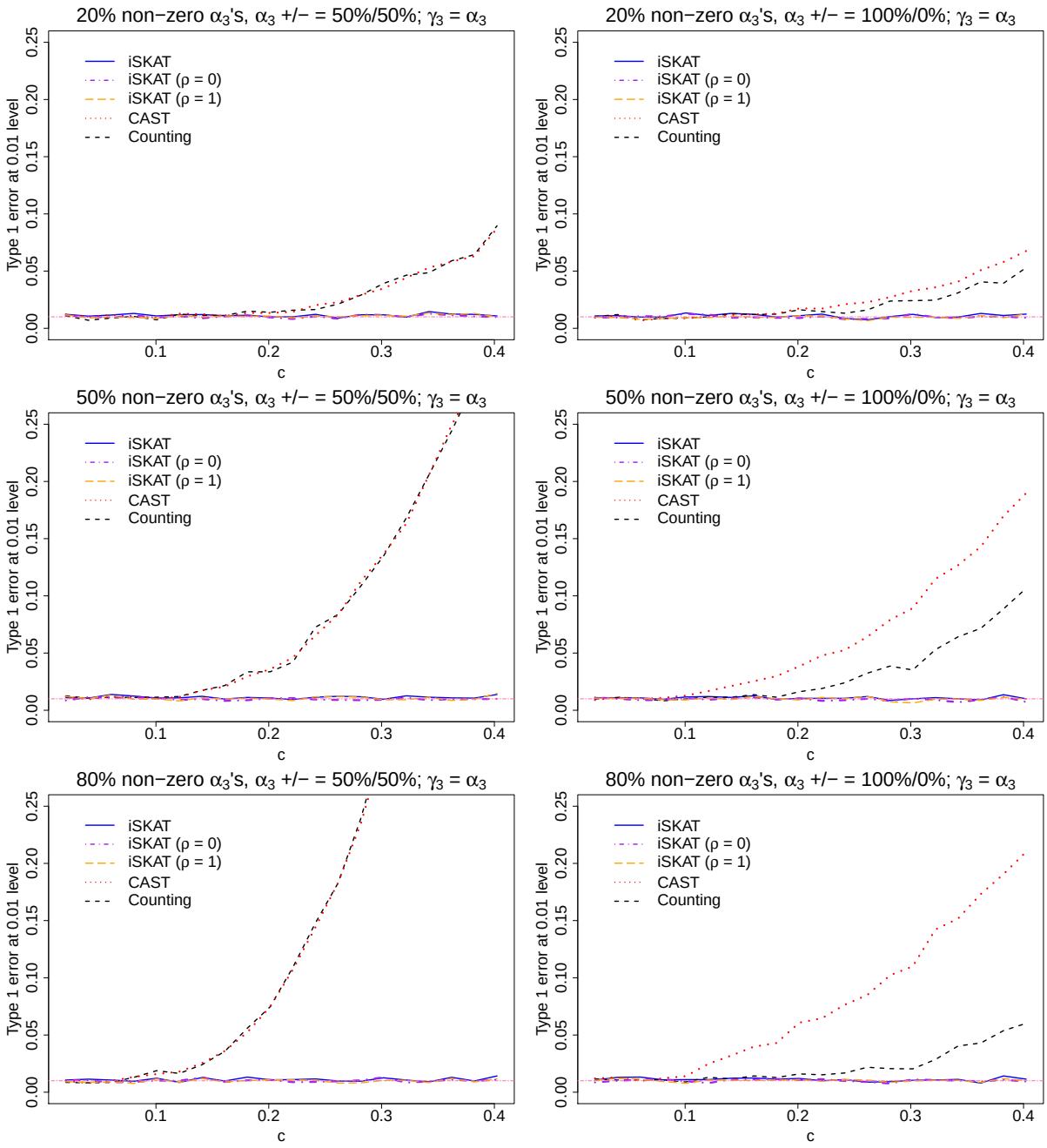
Web Figure 8: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **no main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero γ_{3j}'s; Middle panel - 50% non-zero γ_{3j}'s; Bottom panel - 80% non-zero γ_{3j}'s. Left panel - 50% of γ_{3j}'s are positive; Right panel - 100% of γ_{3j}'s are positive. $\boldsymbol{\alpha}_3 = \mathbf{0}$. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - E$ association increases.** Since there are no main effects for rare variants, all five methods have correct Type 1 error rates.



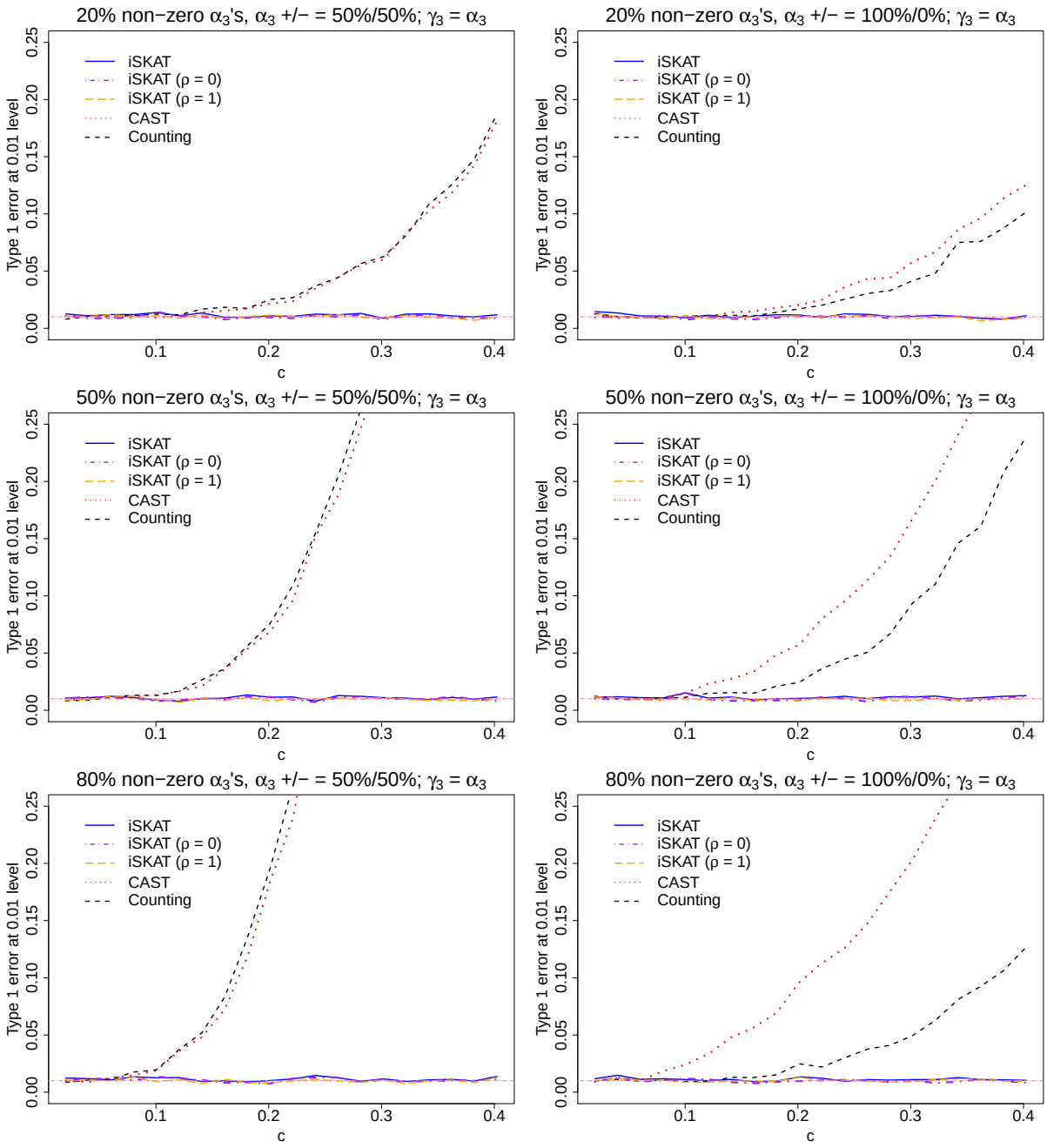
Web Figure 9: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **independent** and there are **main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\boldsymbol{\gamma_3 = 0}$. $\boldsymbol{\alpha_3}$ controls the association between $\boldsymbol{G}$ and Y. **As c increases, the $\boldsymbol{G} - Y$ association increases.** Since there are main effects for rare variants, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates.



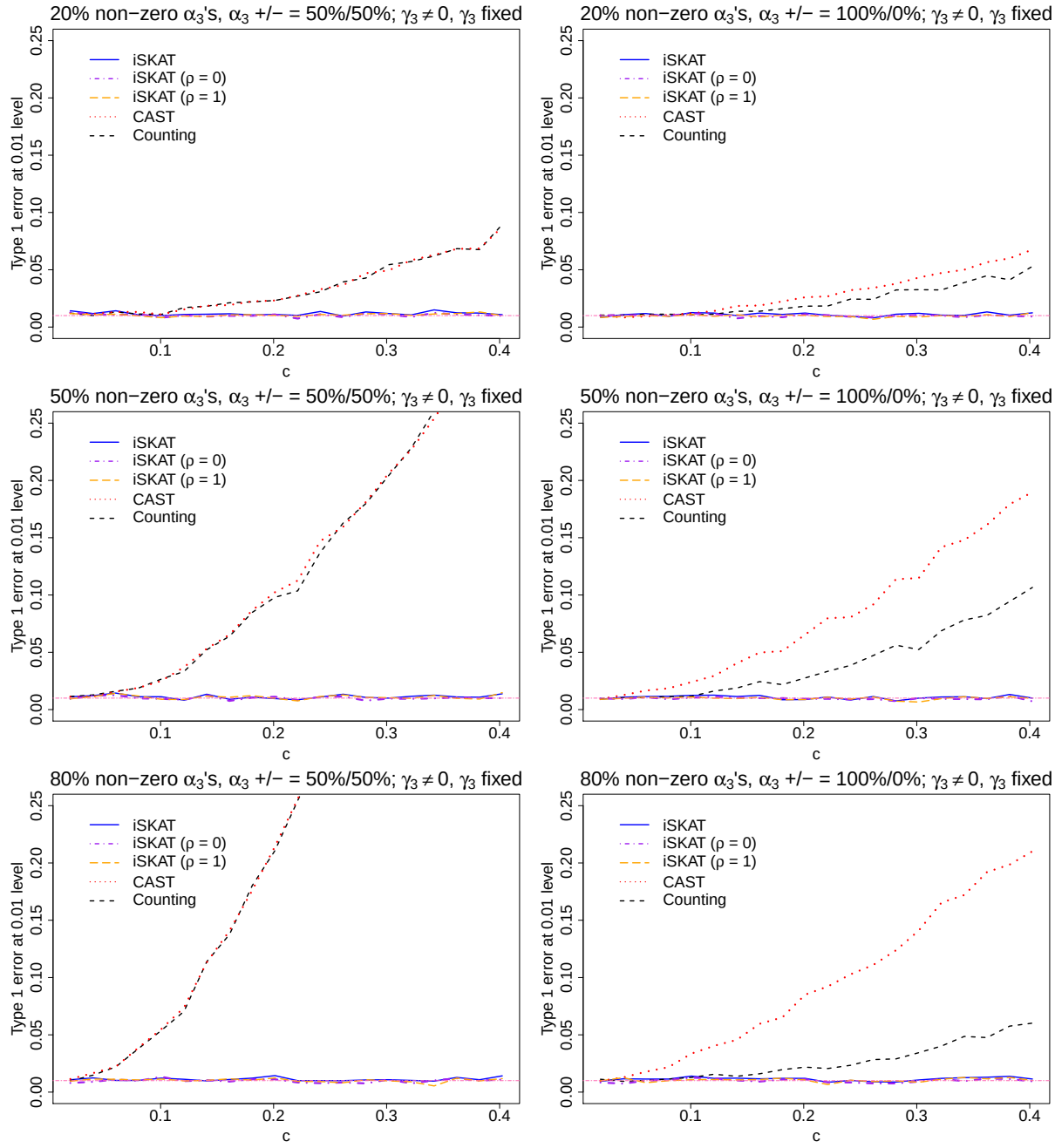
Web Figure 10: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **independent** and there are **main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\boldsymbol{\gamma_3 = 0}$. $\boldsymbol{\alpha_3}$ controls the association between $\boldsymbol{G}$ and Y. **As c increases, the $\boldsymbol{G} - Y$ association increases.** Since there are main effects for rare variants, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates.



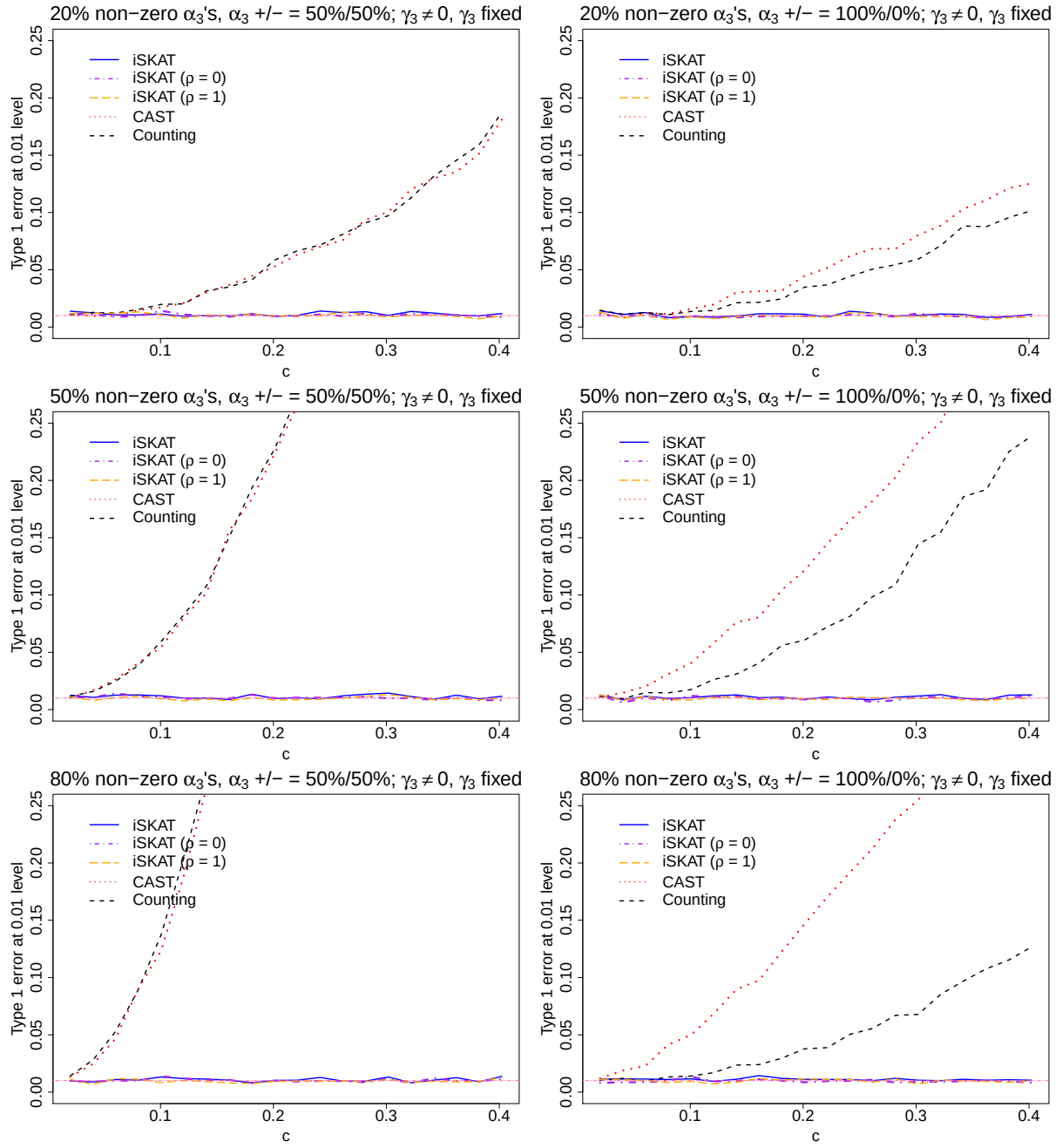
Web Figure 11: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As** c **increases, both** $\boldsymbol{G} - Y$ **and** $\boldsymbol{G} - E$ **associations increase.** Since there are main effects for rare variants, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates.



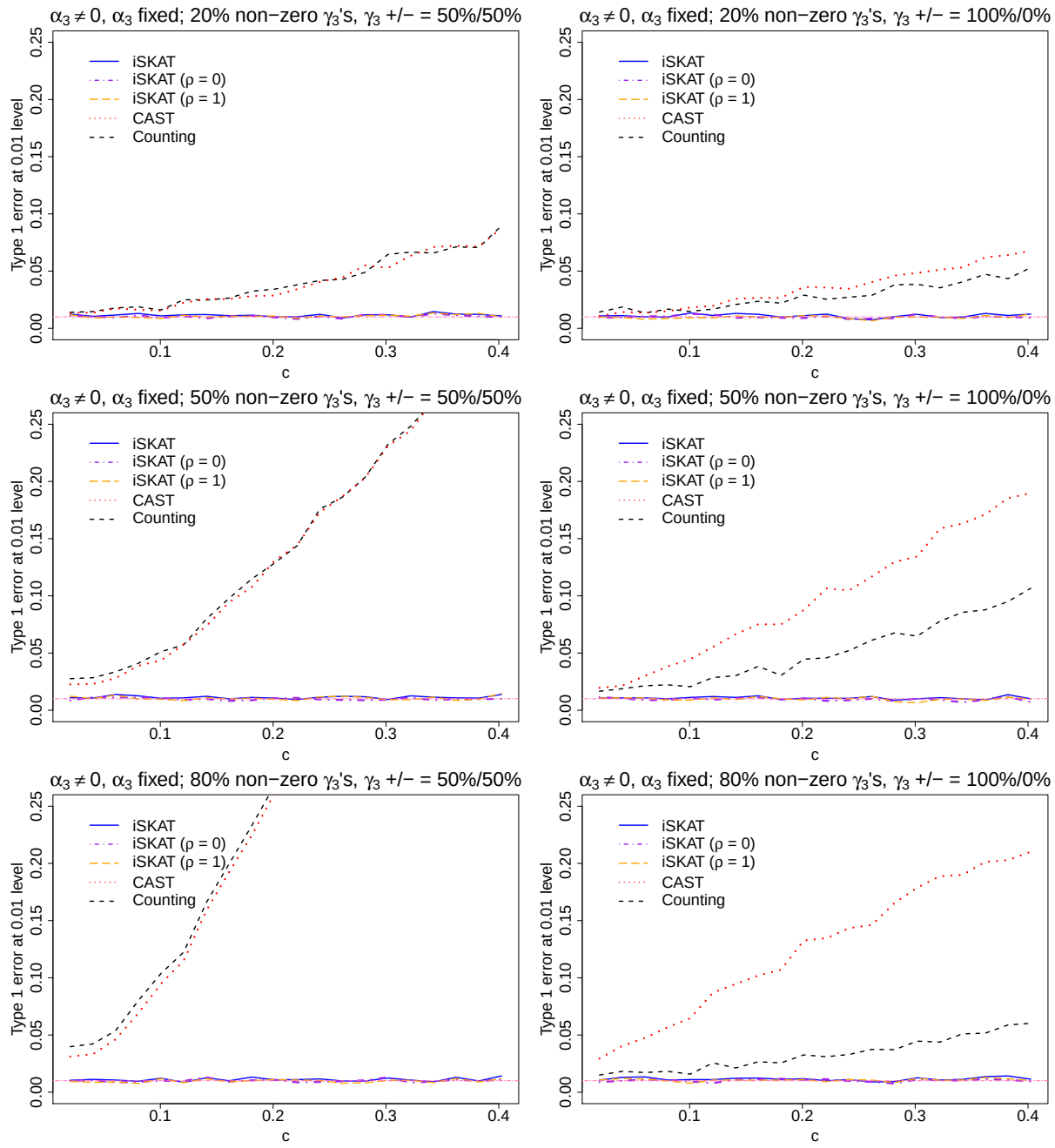
Web Figure 12: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As** c **increases, both** $\boldsymbol{G} - Y$ **and** $\boldsymbol{G} - E$ **associations increase.** Since there are main effects for rare variants, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates.



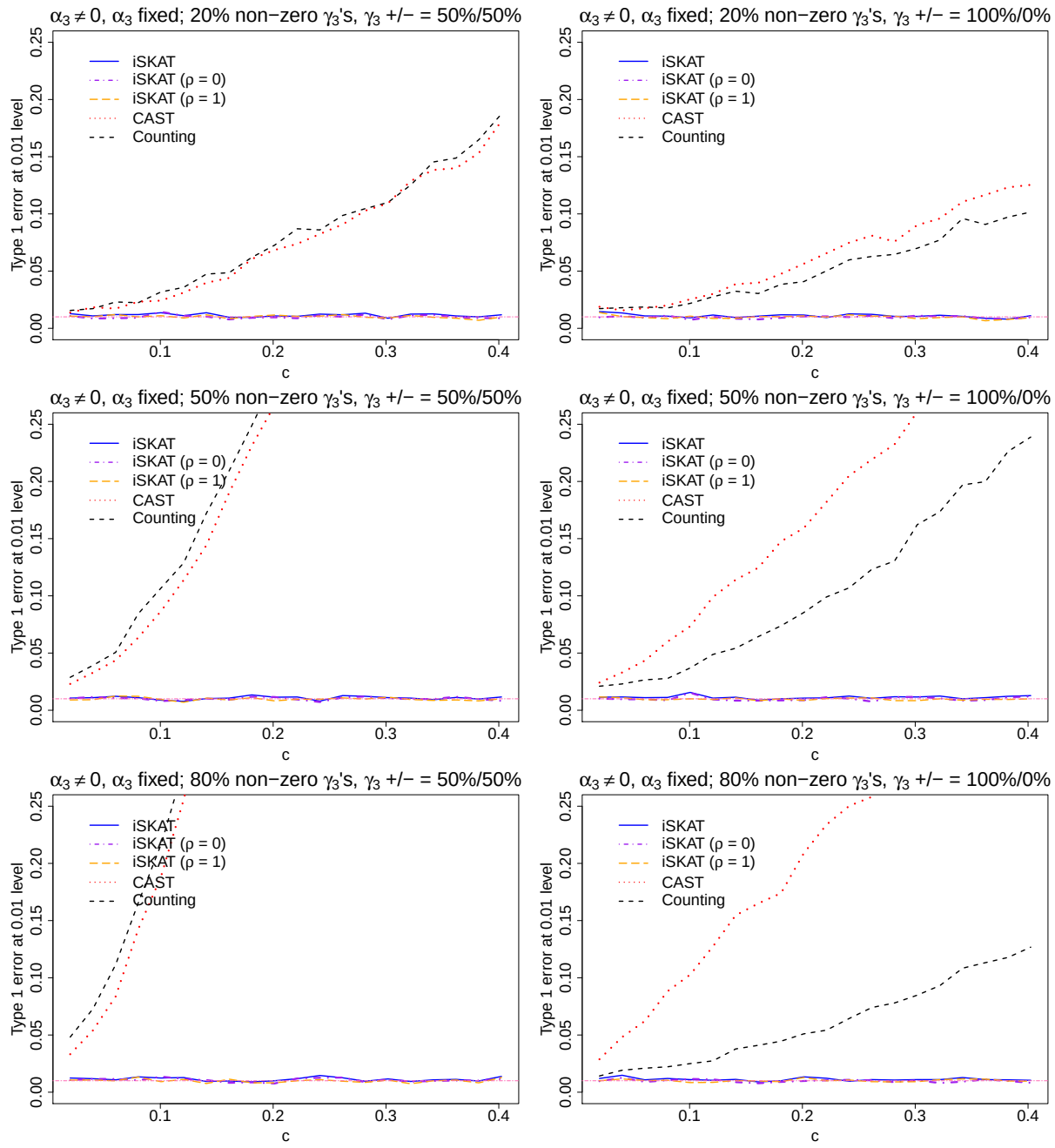
Web Figure 13: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\gamma_{3j} = 0.4\text{sgn}(\alpha_{3j})|\log_{10}\text{MAF}_j|$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - Y$ association increases but $\boldsymbol{G} - E$ association remains fixed.** Since there are main effects for rare variants, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates. Bias of burden tests increase with increasing $\boldsymbol{G} - Y$ association when $\boldsymbol{G}$ and E are dependent.



Web Figure 14: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\gamma_{3j} = 0.4\text{sgn}(\alpha_{3j})|\log_{10}\text{MAF}_j|$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - Y$ association increases but $\boldsymbol{G} - E$ association remains fixed.** Since there are main effects for rare variants, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates. Bias of burden tests increase with increasing $\boldsymbol{G} - Y$ association when $\boldsymbol{G}$ and E are dependent.



Web Figure 15: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero γ_{3j}'s; Middle panel - 50% non-zero γ_{3j}'s; Bottom panel - 80% non-zero γ_{3j}'s. Left panel - 50% of γ_{3j}'s are positive; Right panel - 100% of γ_{3j}'s are positive. $\alpha_{3j} = 0.4\text{sgn}(\gamma_{3j})|\log_{10}\text{MAF}_j|$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - E$ association increases but $\boldsymbol{G} - Y$ association remains fixed.** Since there are main effects for rare variants, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates. Bias of burden tests increase with increasing $\boldsymbol{G} - E$ association when $\boldsymbol{G}$ and Y are associated.

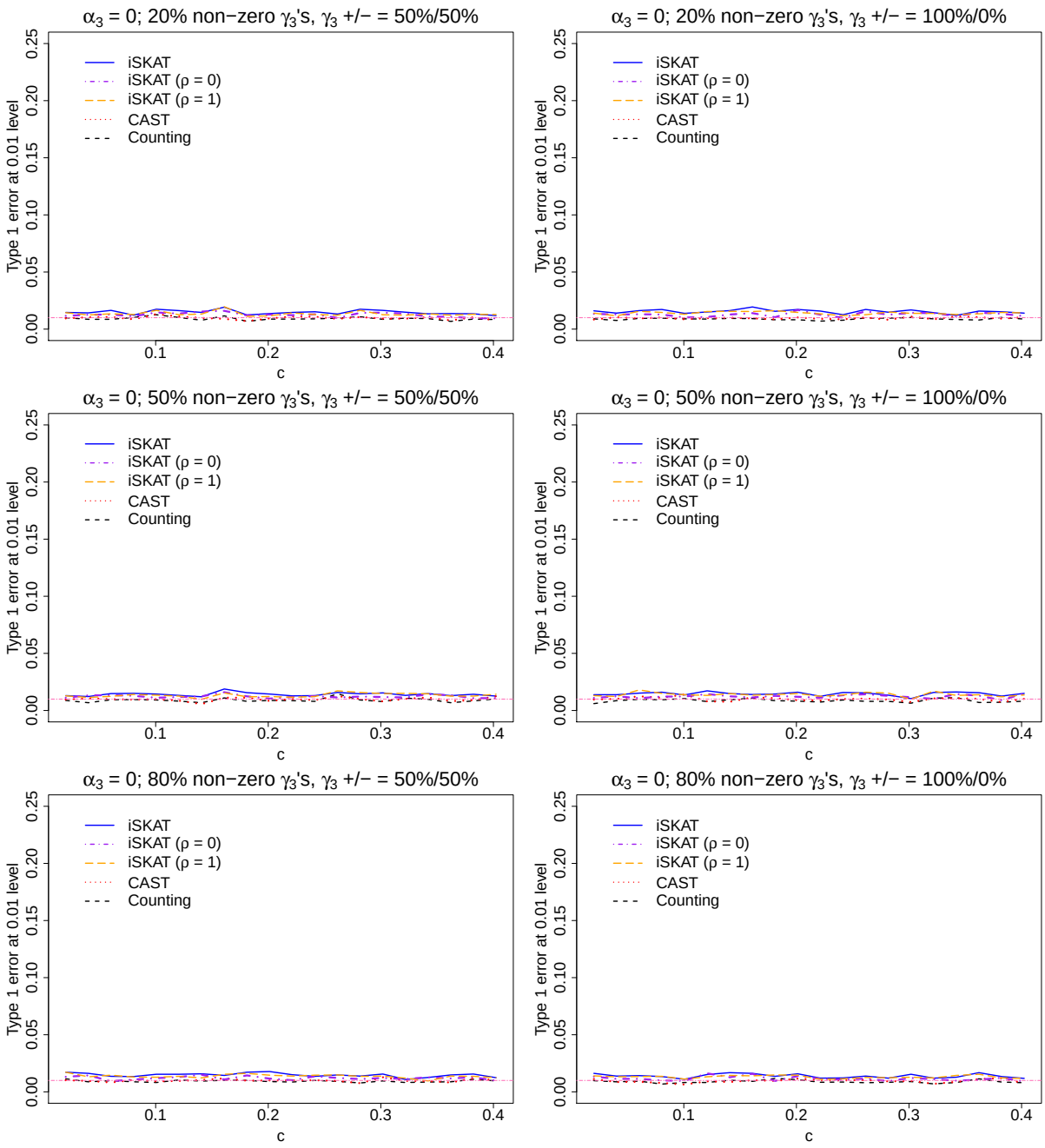


Web Figure 16: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **continuous** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero γ_{3j}'s; Middle panel - 50% non-zero γ_{3j}'s; Bottom panel - 80% non-zero γ_{3j}'s. Left panel - 50% of γ_{3j}'s are positive; Right panel - 100% of γ_{3j}'s are positive. $\alpha_{3j} = 0.4\text{sgn}(\gamma_{3j})|\log_{10}\text{MAF}_j|$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - E$ association increases but $\boldsymbol{G} - Y$ association remains fixed.** Since there are main effects for rare variants, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates. Bias of burden tests increase with increasing $\boldsymbol{G} - E$ association when $\boldsymbol{G}$ and Y are associated.

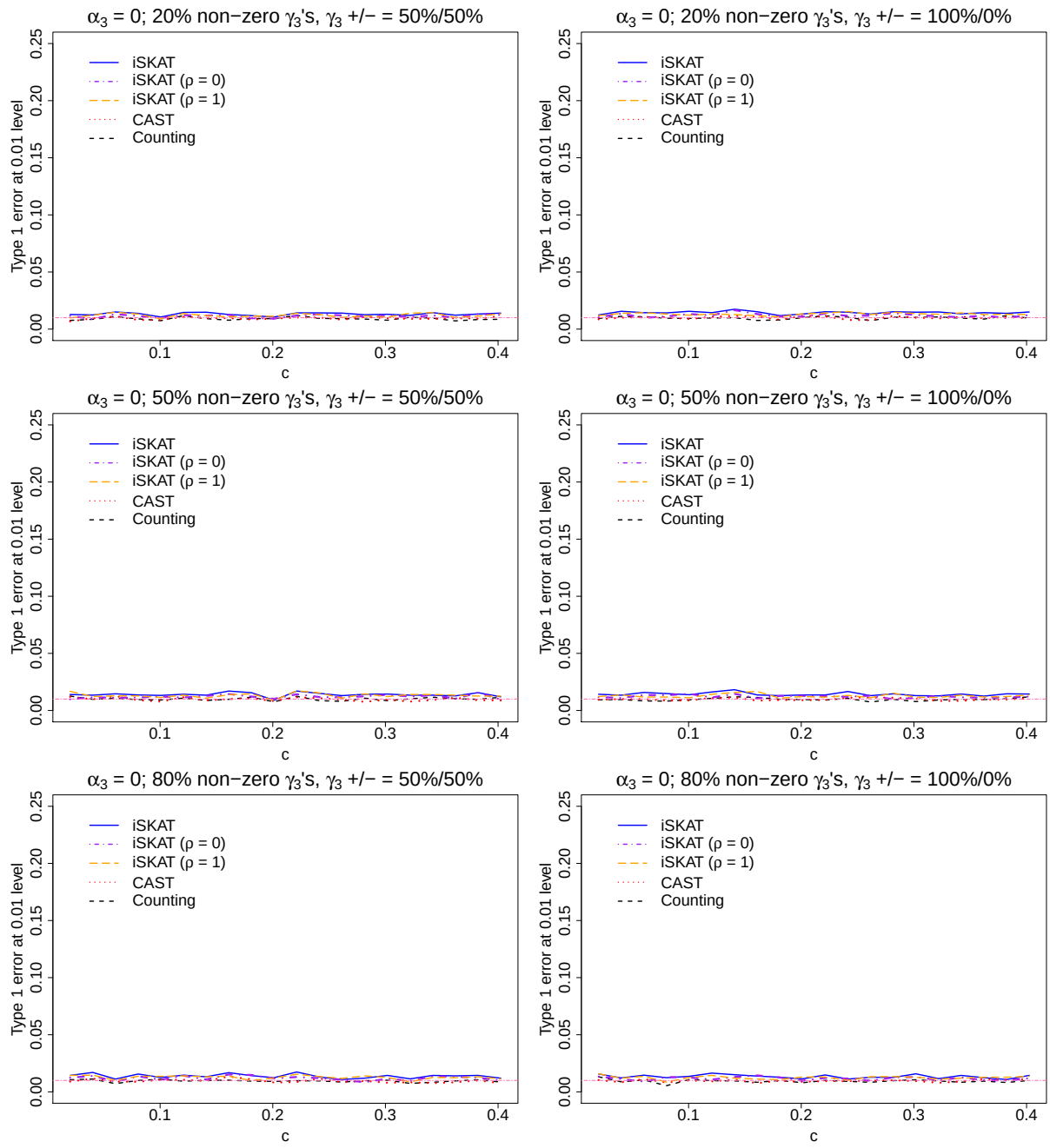
Web Table 3: Empirical Type 1 error rates for **binary** outcome when $\boldsymbol{G}$ and E are **independent** and there are **no main** effects (i.e. $\boldsymbol{\alpha}_3 = \boldsymbol{\gamma}_3 = \mathbf{0}$) for $n = 2000$ (top panel) and $n = 4000$ (bottom panel) respectively. Since there are no main effects for rare variants, all five methods have correct Type 1 error rates.

	α-level	iSKAT	iSKAT ($\rho = 0$)	iSKAT ($\rho = 1$)	CAST	Counting
$n = 2000$	1e-02	1.55e-02	1.31e-02	1.43e-02	1.03e-02	9.77e-03
	1e-03	1.52e-03	1.14e-03	1.51e-03	1.03e-03	8.50e-04
	1e-04	7.00e-05	6.00e-05	1.40e-04	1.10e-04	6.00e-05

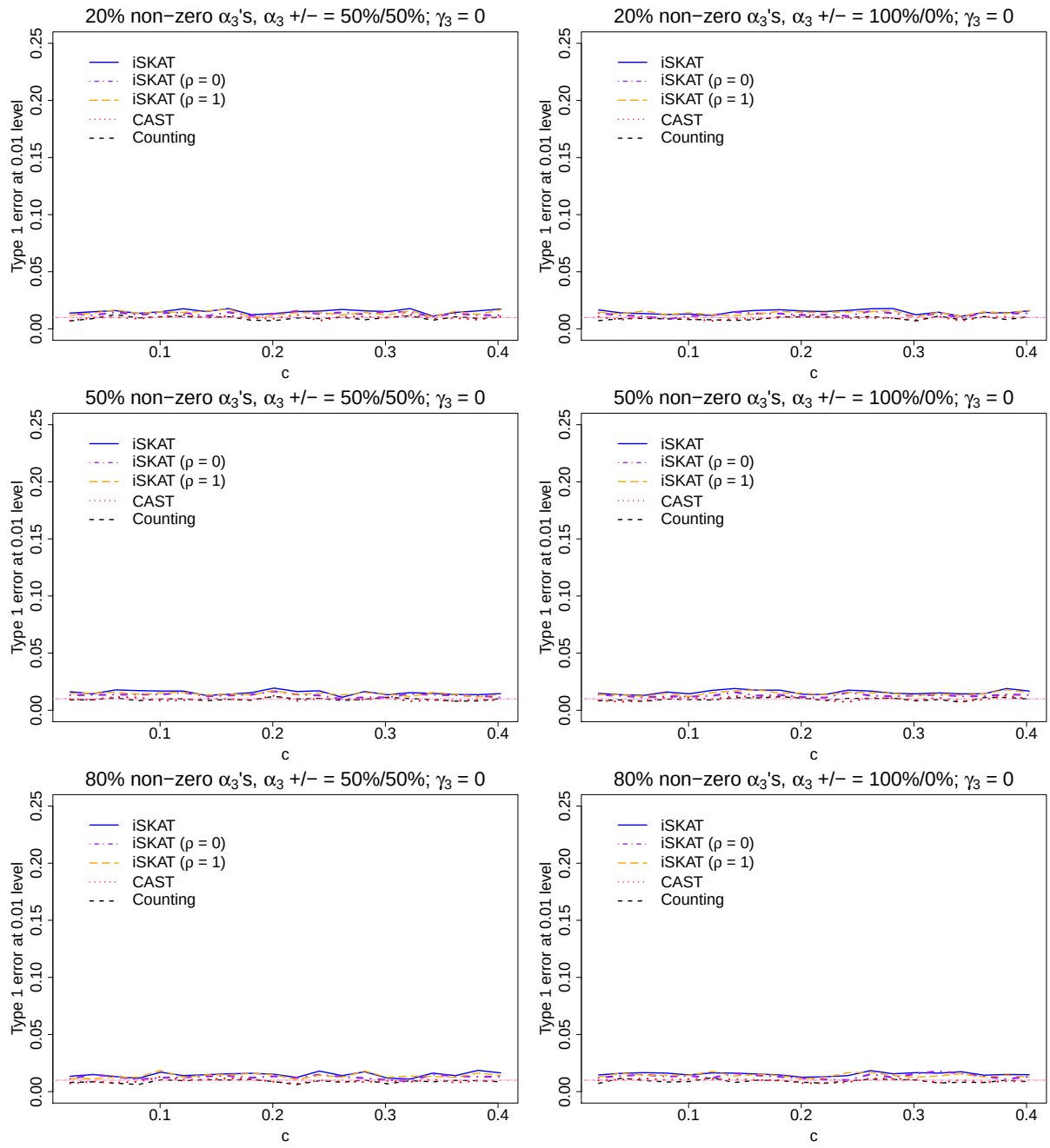
	α-level	iSKAT	iSKAT ($\rho = 0$)	iSKAT ($\rho = 1$)	CAST	Counting
$n = 4000$	1e-02	1.38e-02	1.15e-02	1.30e-02	9.41e-03	9.60e-03
	1e-03	1.43e-03	9.80e-04	1.46e-03	8.70e-04	8.30e-04
	1e-04	7.00e-05	6.00e-05	1.10e-04	5.00e-05	3.00e-05



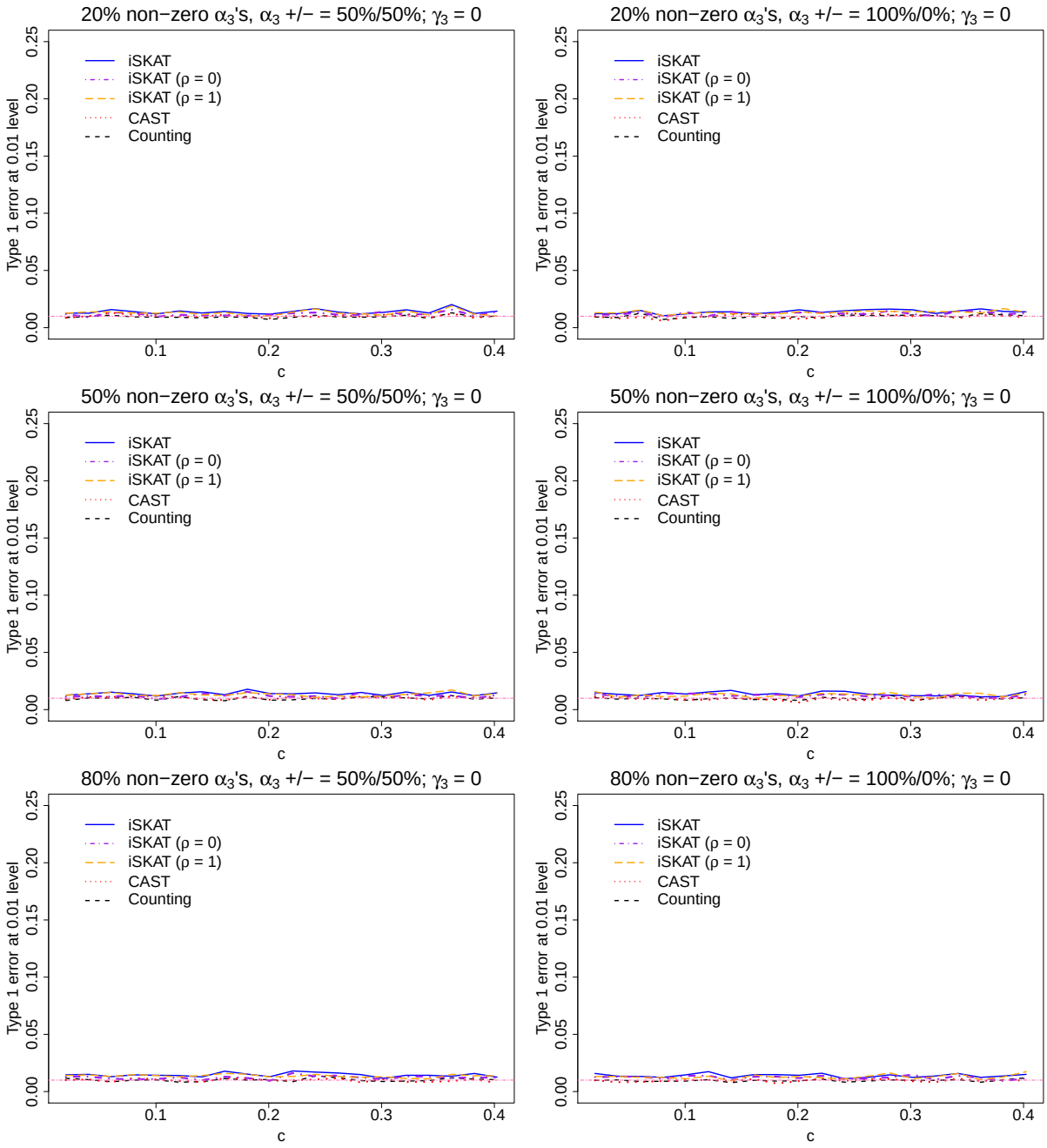
Web Figure 17: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **no main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero γ_{3j}'s; Middle panel - 50% non-zero γ_{3j}'s; Bottom panel - 80% non-zero γ_{3j}'s. Left panel - 50% of γ_{3j}'s are positive; Right panel - 100% of γ_{3j}'s are positive. $\boldsymbol{\alpha}_3 = \mathbf{0}$. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - E$ association increases.** Since there are no main effects for rare variants, all five methods have correct Type 1 error rates.



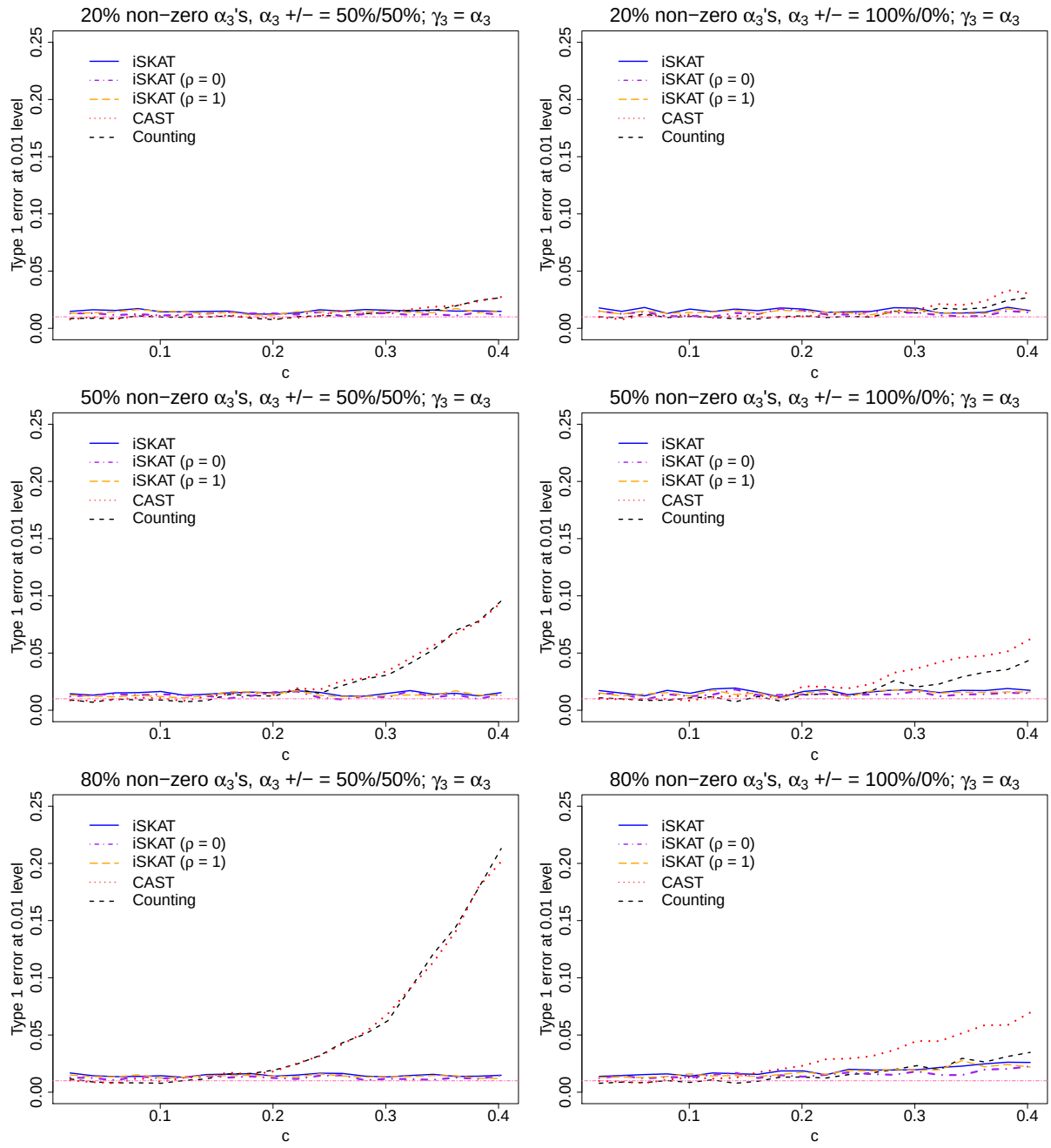
Web Figure 18: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **no main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero γ_{3j}'s; Middle panel - 50% non-zero γ_{3j}'s; Bottom panel - 80% non-zero γ_{3j}'s. Left panel - 50% of γ_{3j}'s are positive; Right panel - 100% of γ_{3j}'s are positive. $\boldsymbol{\alpha}_3 = \mathbf{0}$. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As** c **increases, the** $\boldsymbol{G} - E$ **association increases.** Since there are no main effects for rare variants, all five methods have correct Type 1 error rates.



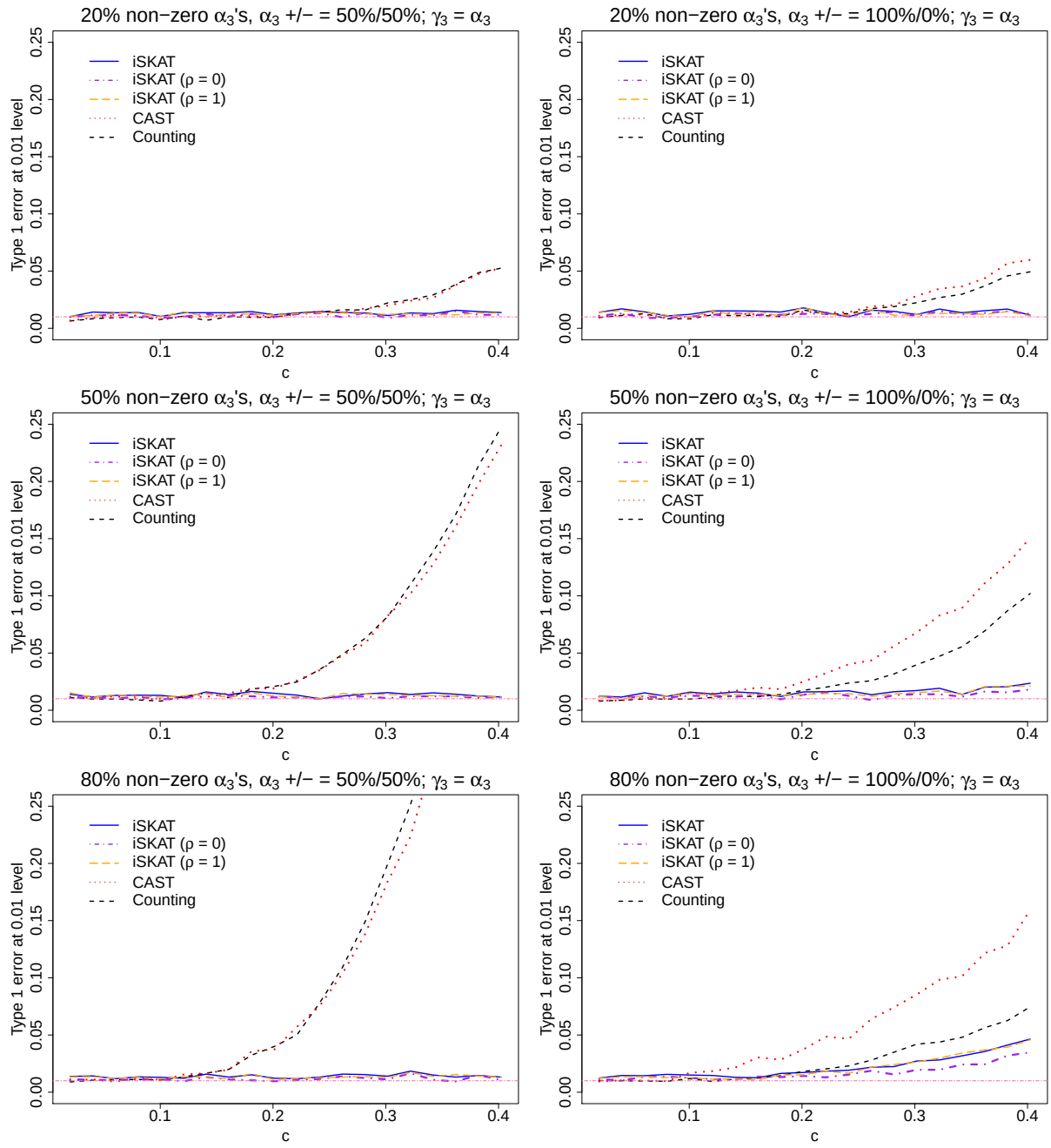
Web Figure 19: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **independent** and there are **main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\boldsymbol{\gamma_3 = 0}$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. **As c increases, the $\boldsymbol{G} - Y$ association increases.** Since $\boldsymbol{G}$ and E are independent and outcome is binary, even though there are main effects for rare variants, all five methods have correct Type 1 error rates.



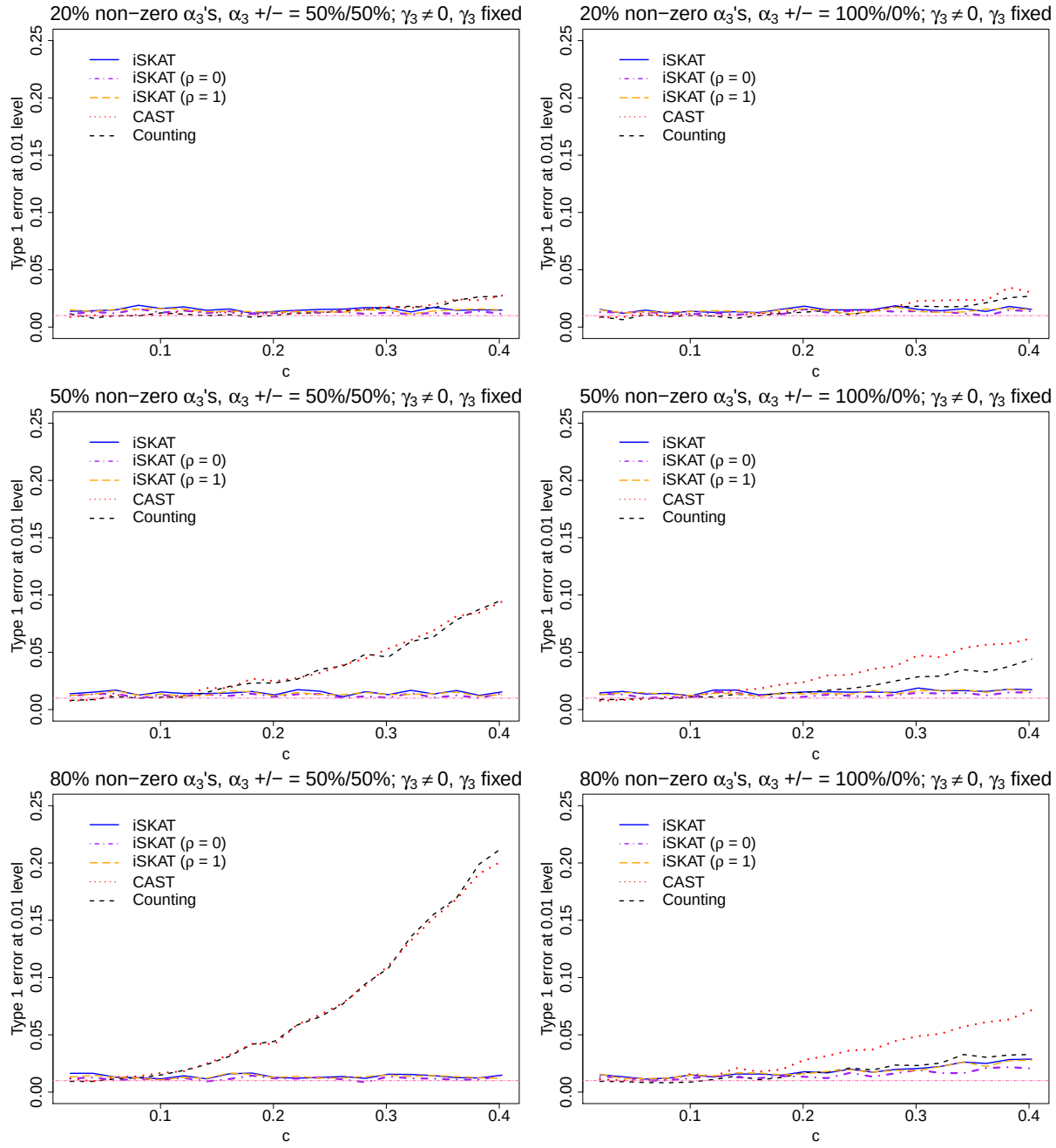
Web Figure 20: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **independent** and there are **main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\boldsymbol{\gamma_3 = 0}$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. **As c increases, the $\boldsymbol{G} - Y$ association increases.** Since $\boldsymbol{G}$ and E are independent and outcome is binary, even though there are main effects for rare variants, all five methods have correct Type 1 error rates.



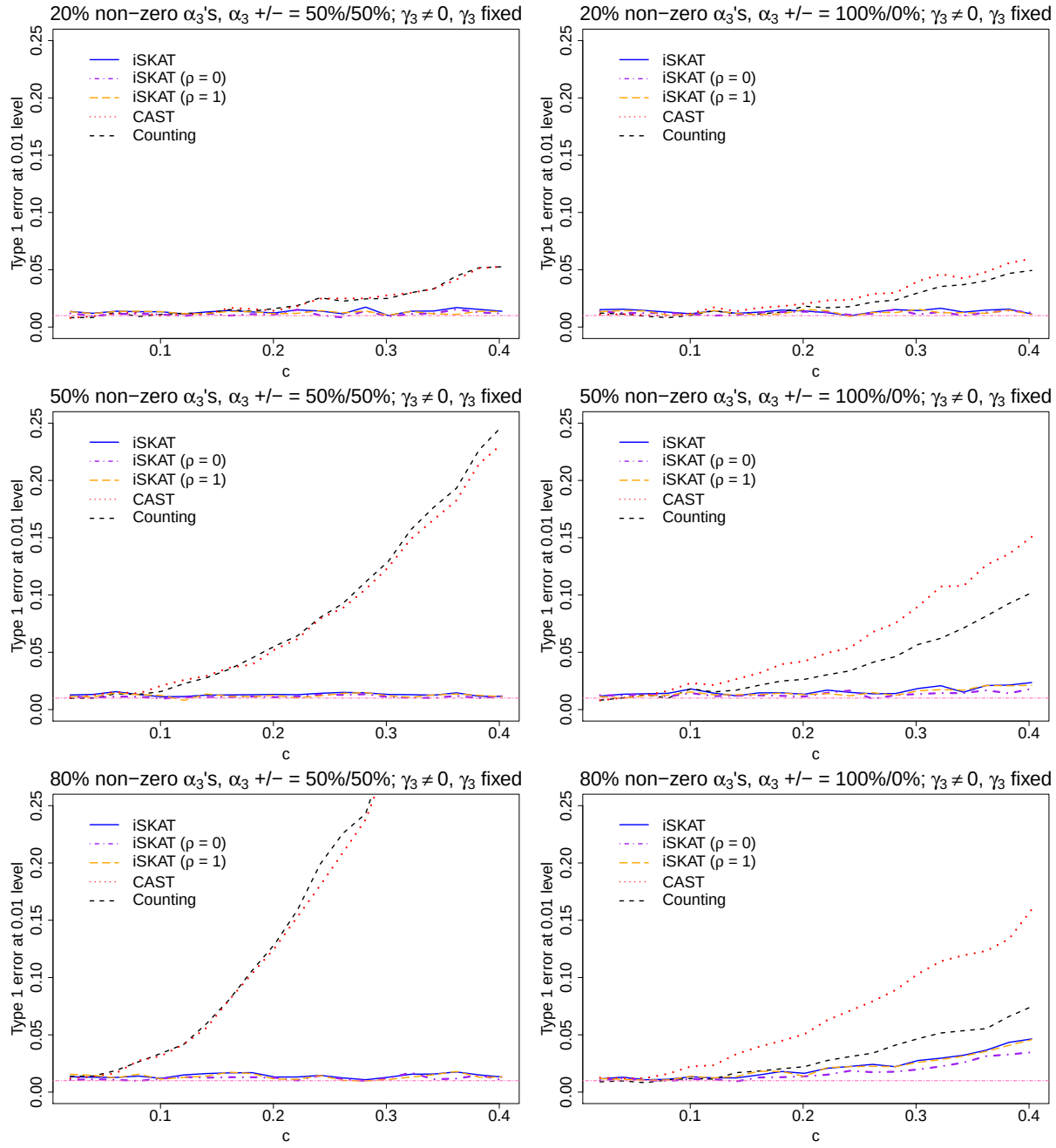
Web Figure 21: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, both $\boldsymbol{G} - Y$ and $\boldsymbol{G} - E$ associations increase.** Since there are main effects for rare variants and $\boldsymbol{G}$ and E are dependent, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates.



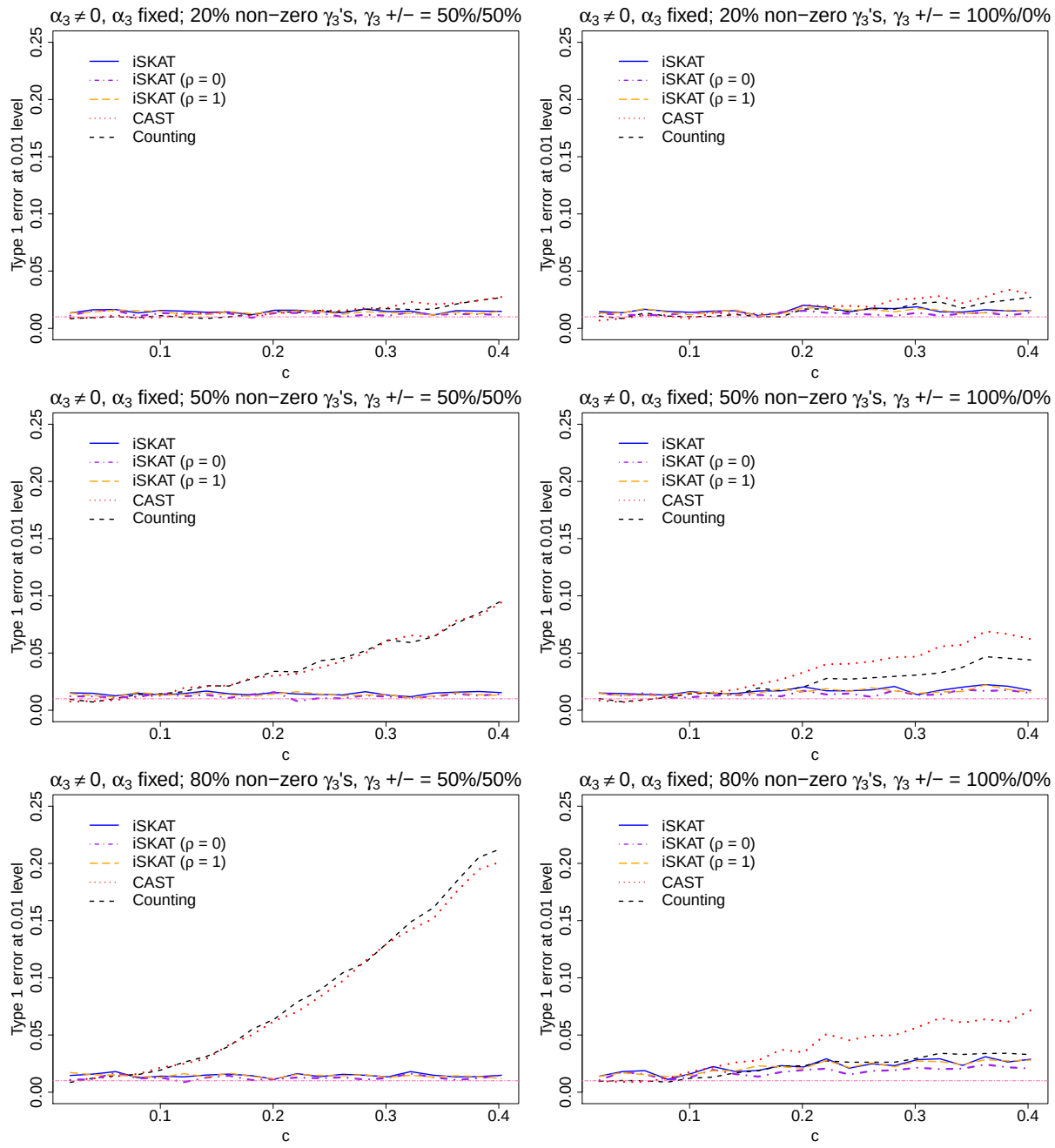
Web Figure 22: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, both $\boldsymbol{G} - Y$ and $\boldsymbol{G} - E$ associations increase.** Since there are main effects for rare variants and $\boldsymbol{G}$ and E are dependent, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates.



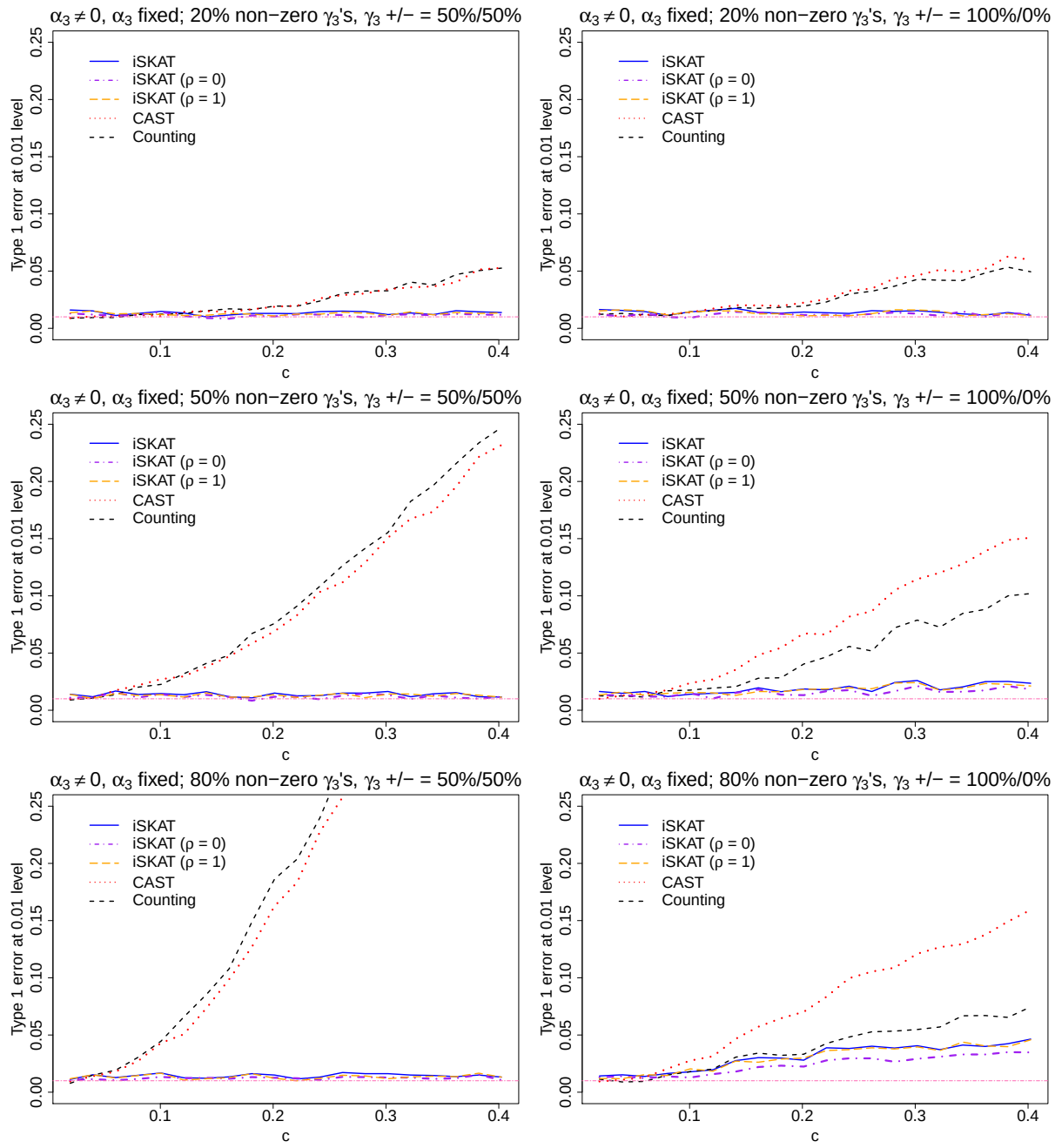
Web Figure 23: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\gamma_{3j} = 0.4\text{sgn}(\alpha_{3j})|\log_{10}\text{MAF}_j|$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - Y$ association increases but $\boldsymbol{G} - E$ association remains fixed.** Since there are main effects for rare variants and $\boldsymbol{G}$ and E are dependent, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates. Bias of burden tests increase with increasing $\boldsymbol{G} - Y$ association when $\boldsymbol{G}$ and E are dependent.



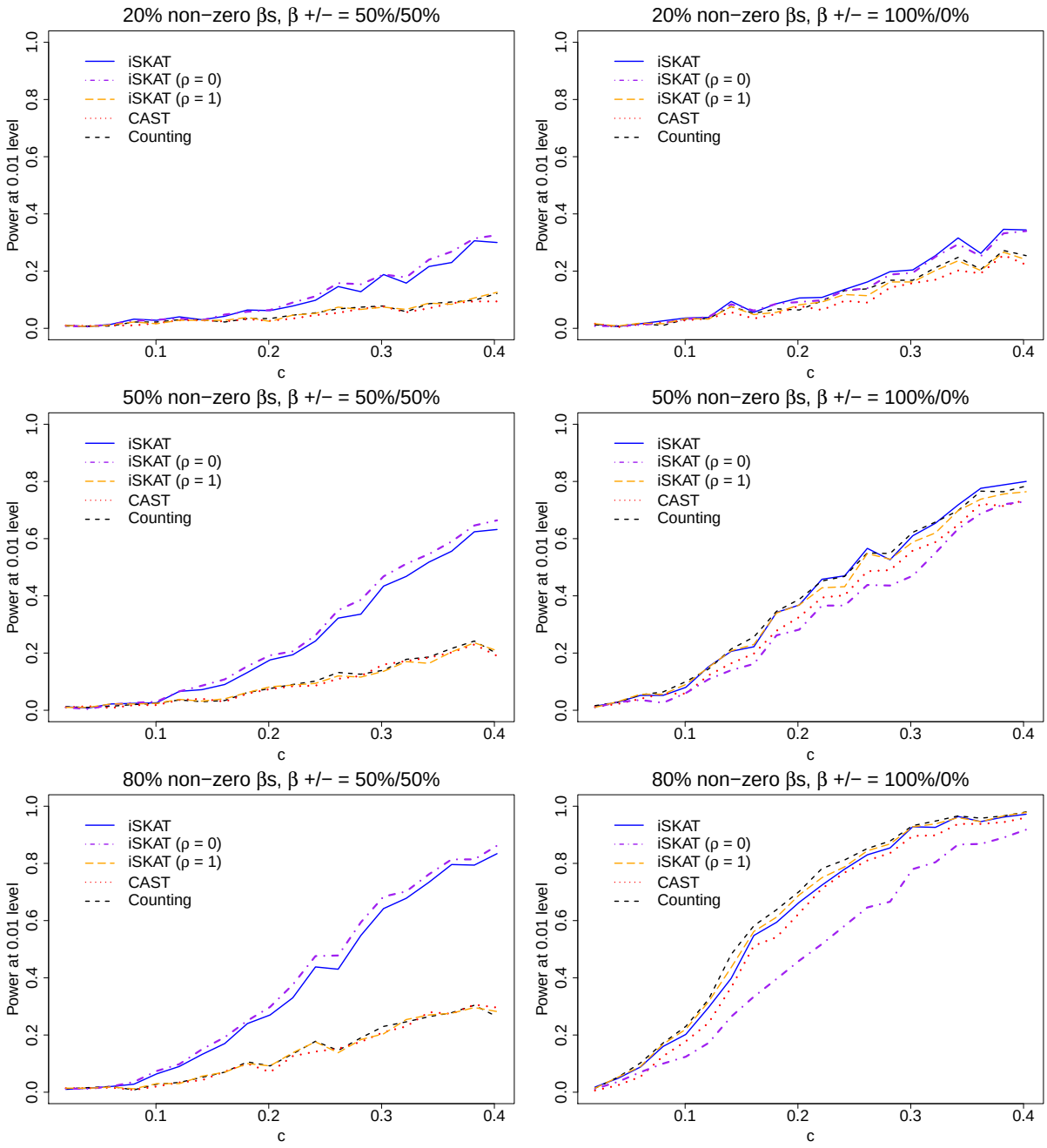
Web Figure 24: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero α_{3j}'s; Middle panel - 50% non-zero α_{3j}'s; Bottom panel - 80% non-zero α_{3j}'s. Left panel - 50% of α_{3j}'s are positive; Right panel - 100% of α_{3j}'s are positive. $\gamma_{3j} = 0.4\text{sgn}(\alpha_{3j})|\log_{10}\text{MAF}_j|$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - Y$ association increases but $\boldsymbol{G} - E$ association remains fixed.** Since there are main effects for rare variants and $\boldsymbol{G}$ and E are dependent, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates. Bias of burden tests increase with increasing $\boldsymbol{G} - Y$ association when $\boldsymbol{G}$ and E are dependent.



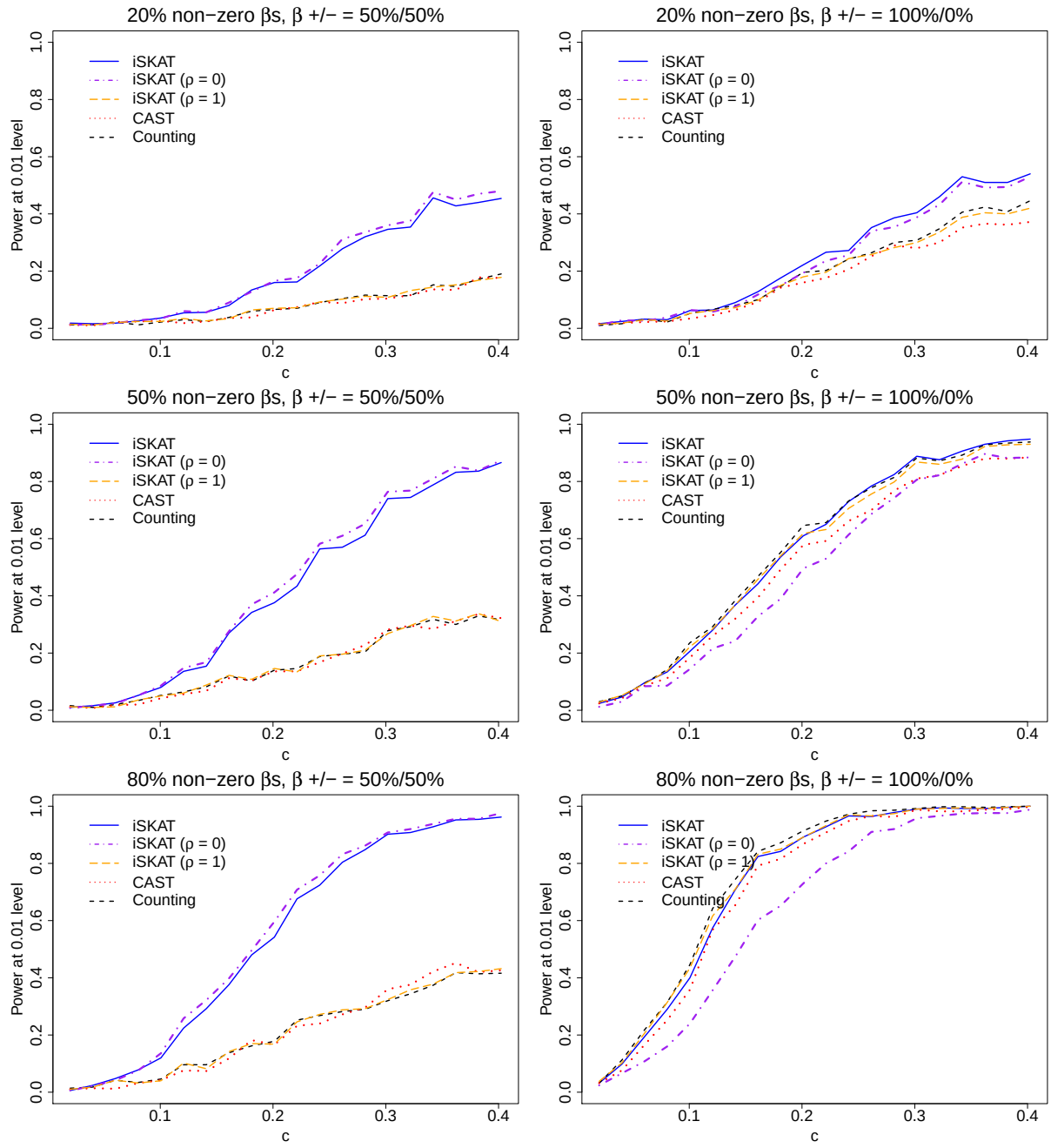
Web Figure 25: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 2000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero γ_{3j}'s; Middle panel - 50% non-zero γ_{3j}'s; Bottom panel - 80% non-zero γ_{3j}'s. Left panel - 50% of γ_{3j}'s are positive; Right panel - 100% of γ_{3j}'s are positive. $\alpha_{3j} = 0.4\text{sgn}(\gamma_{3j})|\log_{10}\text{MAF}_j|$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - E$ association increases but $\boldsymbol{G} - Y$ association remains fixed.** Since there are main effects for rare variants and $\boldsymbol{G}$ and E are dependent, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates. Bias of burden tests increase with increasing $\boldsymbol{G} - E$ association when $\boldsymbol{G}$ and Y are associated.



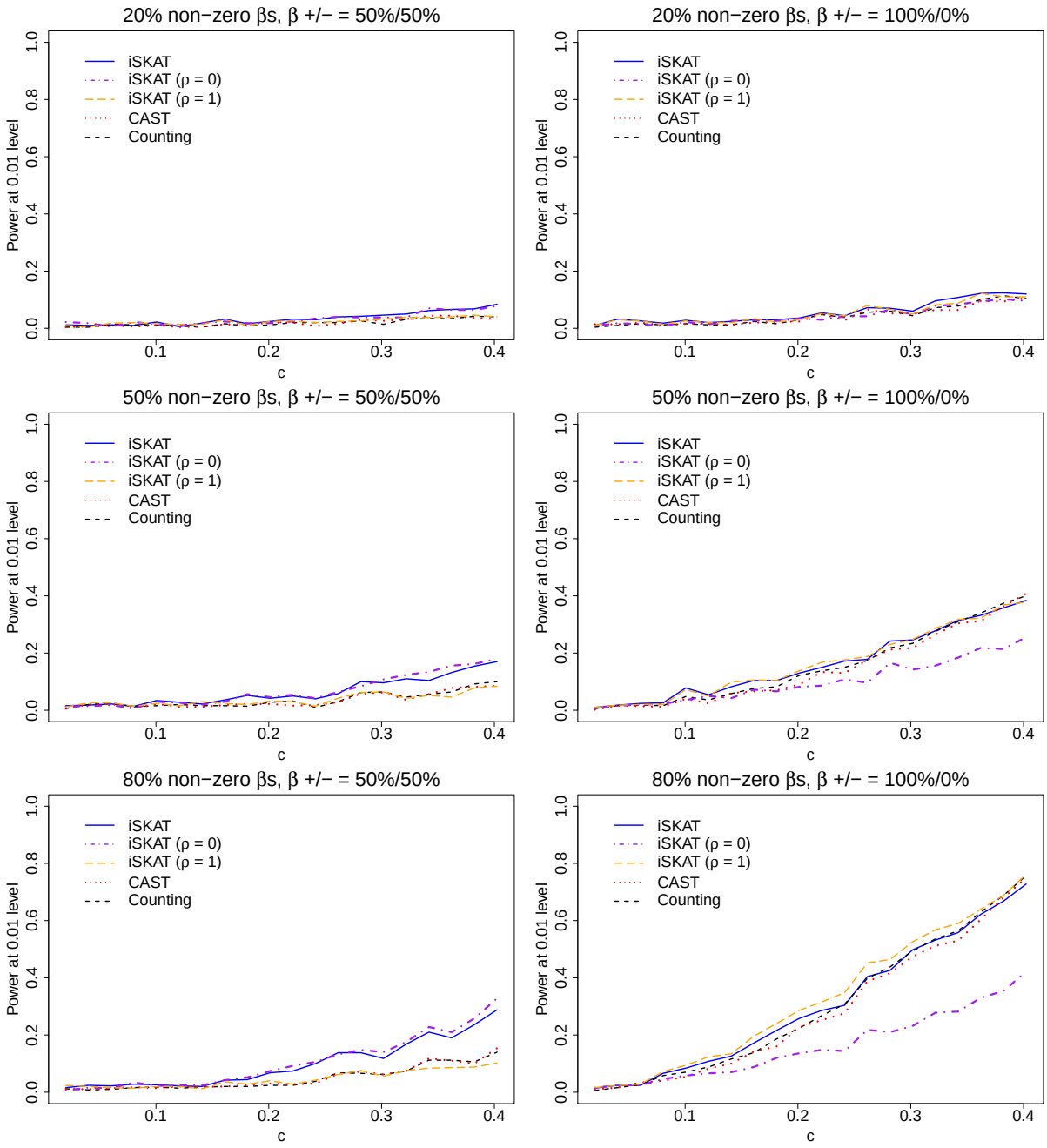
Web Figure 26: Empirical Type 1 error rates at $\alpha = 0.01$ nominal level for **binary** outcome when $\boldsymbol{G}$ and E are **dependent** and there are **main** effects for $n = 4000$ - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero γ_{3j}'s; Middle panel - 50% non-zero γ_{3j}'s; Bottom panel - 80% non-zero γ_{3j}'s. Left panel - 50% of γ_{3j}'s are positive; Right panel - 100% of γ_{3j}'s are positive. $\alpha_{3j} = 0.4\text{sgn}(\gamma_{3j})|\log_{10}\text{MAF}_j|$. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y. $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E. **As c increases, the $\boldsymbol{G} - E$ association increases but $\boldsymbol{G} - Y$ association remains fixed.** Since there are main effects for rare variants and $\boldsymbol{G}$ and E are dependent, iSKAT gives a correct Type 1 error rate but burden tests can have inflated Type 1 error rates. Bias of burden tests increase with increasing $\boldsymbol{G} - E$ association when $\boldsymbol{G}$ and Y are associated.



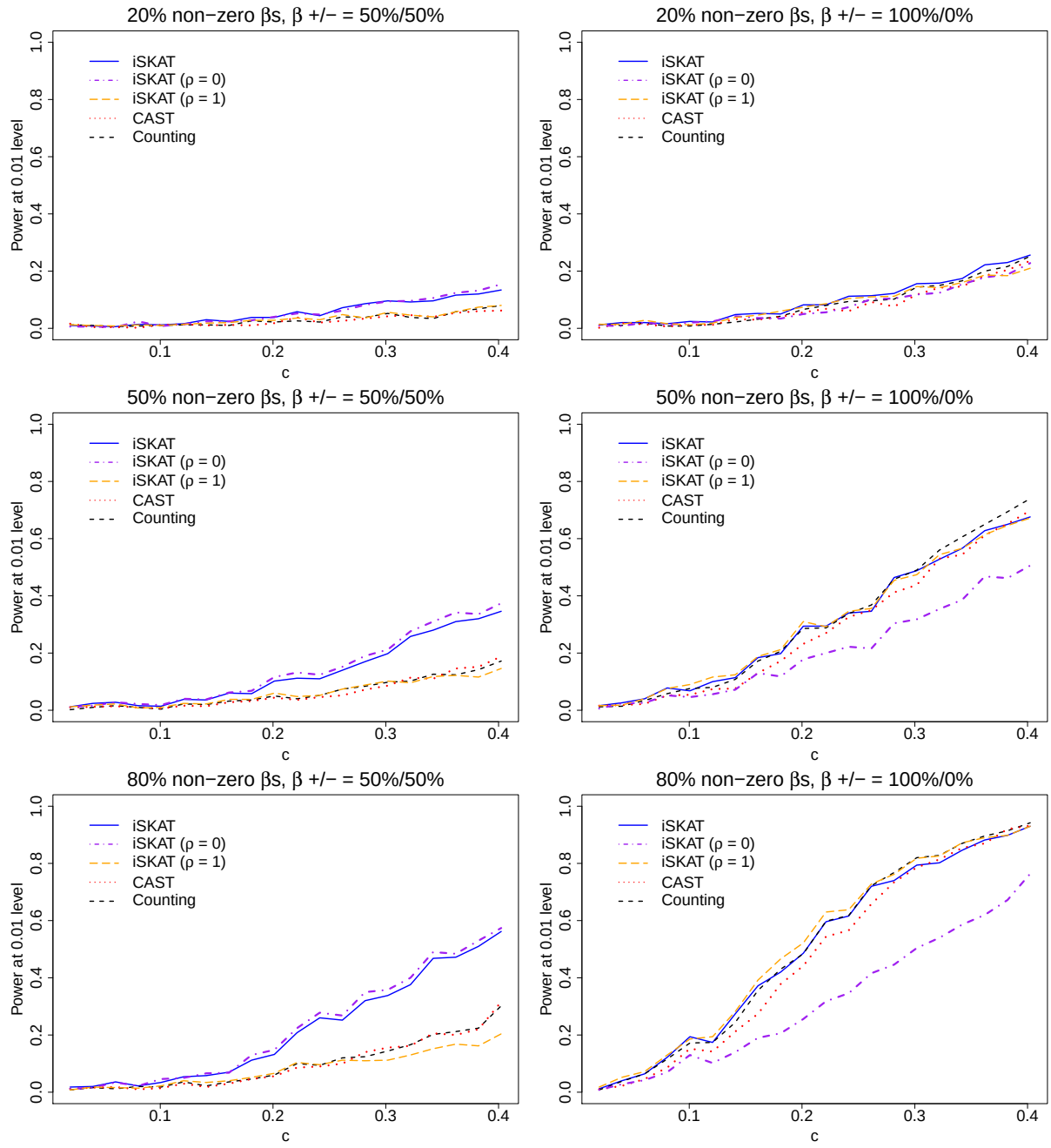
Web Figure 27: Empirical power curves at $n = 2000$ for $\alpha = 0.01$ level of significance for **continuous** outcome when $\boldsymbol{G}$ and E are **independent** and there are **no main** effects (i.e. $\boldsymbol{\alpha}_3 = \boldsymbol{\gamma}_3 = \mathbf{0}$) - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero β_j's; Middle panel - 50% non-zero β_j's; Bottom panel - 80% non-zero β_j's. Left panel - 50% of β_j's are positive; Right panel - 100% of β_j's are positive.



Web Figure 28: Empirical power curves at $n = 4000$ for $\alpha = 0.01$ level of significance for **continuous** outcome when $\boldsymbol{G}$ and E are **independent** and there are **no main** effects (i.e. $\boldsymbol{\alpha}_3 = \boldsymbol{\gamma}_3 = \mathbf{0}$) - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero β_j's; Middle panel - 50% non-zero β_j's; Bottom panel - 80% non-zero β_j's. Left panel - 50% of β_j's are positive; Right panel - 100% of β_j's are positive.



Web Figure 29: Empirical power curves at $n = 2000$ for $\alpha = 0.01$ level of significance for **binary** outcome when $\boldsymbol{G}$ and E are **independent** and there are **no main** effects (i.e. $\boldsymbol{\alpha}_3 = \boldsymbol{\gamma}_3 = \mathbf{0}$) - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero β_j's; Middle panel - 50% non-zero β_j's; Bottom panel - 80% non-zero β_j's. Left panel - 50% of β_j's are positive; Right panel - 100% of β_j's are positive.



Web Figure 30: Empirical power curves at $n = 4000$ for $\alpha = 0.01$ level of significance for **binary** outcome when $\boldsymbol{G}$ and E are **independent** and there are **no main** effects (i.e. $\boldsymbol{\alpha}_3 = \boldsymbol{\gamma}_3 = \mathbf{0}$) - iSKAT (solid line), iSKAT with $\rho = 0$ (dashed-and-dotted line), iSKAT with $\rho = 1$ (long dashed line), CAST (dotted line) and Counting (short dashed line). Top panel - 20% non-zero β_j's; Middle panel - 50% non-zero β_j's; Bottom panel - 80% non-zero β_j's. Left panel - 50% of β_j's are positive; Right panel - 100% of β_j's are positive.

Web Appendix 5. Additional Data Analysis on the CoLaus Dataset

For the data analysis on the CoLaus resequencing dataset, we applied a log10 transformation to the plasma adiponectin levels and extreme values of adiponectin levels were set to the boundary value. This was done to improve normality and lessen the impact of outliers. Web Figure 31 shows normal quantile-quantile plots, boxplots and histograms of adiponectin levels both before and after transformation.

The minor allele frequencies (MAF) and missing rates of each of the 11 rare variants within the exon of the adiponectin gene in the CoLaus resequencing dataset are provided in Web Table 6. Linkage disequilibrium (LD) measures of these 11 rare variants with chr3:188053586 (MAF=0.138) are given in Web Table 11.

The results from applying SKAT-O to assess the main effects of the 11 rare variants on adiponectin levels are given in Web Table 7. The results from applying iSKAT to assess the rare variants by alcohol interactions on adiponectin levels are given in Web Table 8.

The estimated rare variants main effects from ridge regression and unpenalized regression (ridge parameter $\lambda = 0$) are given in Web Tables 9 and 10 respectively.

In main manuscript Table 2, we report association analysis results where the missing genotypes were imputed with homozygote of the major allele. Web Table 12 gives association analysis results when the missing genotypes were imputed with the mean. Web Table 13 gives association analysis results when the missing genotypes were imputed with homozygote of the major allele, but the six singletons were excluded from the analysis. All three sets of results are similar.

Web Table 4: Empirical Type 1 error rates from iSKAT (null model estimated without penalty i.e. ridge parameter $\lambda = 0$) and iSKAT (null model estimated with penalty i.e. ridge parameter $\lambda > 0$) are given in second and third columns respectively, for continuous outcome when there are main effects ($\boldsymbol{\alpha}_3 \neq \mathbf{0}$), $\boldsymbol{G}$ and E are dependent ($\boldsymbol{\gamma}_3 \neq \mathbf{0}$), $n = 2000$, under the null hypothesis ($\boldsymbol{\beta} = \mathbf{0}$). $\boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$. We consider the cases when 20% (first and second panels), 50% (third and fourth panels) or 80% (fifth and sixth panels) of the α_{3j}'s are non -zero; and when 50% (first, third and fifth panels) or 100% (second, fourth and sixth panels) of α_{3j}'s are positive. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y (main rare variants effects on Y). $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E (main rare variants effects on E). $\boldsymbol{\beta}$ reflects the rare variants by environment interactions. For 71% of the 10^5 simulations, iSKAT with null model estimated without penalty ($\lambda = 0$) did not converge. Here, we only compare the two when iSKAT with null model estimated without penalty ($\lambda = 0$) converged. Type 1 error rates from both approaches are similar for the remaining 29% of simulations.

20% non-zero α_{3j}'s, 50% of α_{3j}'s are positive, $\boldsymbol{\beta} = \mathbf{0}$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	1.15e-02	1.15e-02
1e-03	1.21e-03	1.21e-03

20% non-zero α_{3j}'s, 100% of α_{3j}'s are positive, $\boldsymbol{\beta} = \mathbf{0}$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	1.03e-02	1.04e-02
1e-03	1.00e-03	1.07e-03

50% non-zero α_{3j}'s, 50% of α_{3j}'s are positive, $\boldsymbol{\beta} = \mathbf{0}$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	1.09e-02	1.10e-02
1e-03	1.07e-03	1.07e-03

50% non-zero α_{3j}'s, 100% of α_{3j}'s are positive, $\boldsymbol{\beta} = \mathbf{0}$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	1.06e-02	1.05e-02
1e-03	9.30e-04	9.65e-04

80% non-zero α_{3j}'s, 50% of α_{3j}'s are positive, $\boldsymbol{\beta} = \mathbf{0}$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	1.11e-02	1.11e-02
1e-03	1.17e-03	1.14e-03

80% non-zero α_{3j}'s, 100% of α_{3j}'s are positive, $\boldsymbol{\beta} = \mathbf{0}$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	1.04e-02	1.04e-02
1e-03	1.41e-03	1.38e-03

Web Table 5: Empirical power from iSKAT (null model estimated without penalty i.e. ridge parameter $\lambda = 0$) and iSKAT (null model estimated with penalty i.e. ridge parameter $\lambda > 0$) are given in second and third columns respectively, for continuous outcome when there are main effects ($\boldsymbol{\alpha}_3 \neq \mathbf{0}$), $\boldsymbol{G}$ and E are dependent ($\boldsymbol{\gamma}_3 \neq \mathbf{0}$), $n = 2000$, under the alternative hypothesis ($\boldsymbol{\beta} \neq \mathbf{0}$). $\boldsymbol{\beta} = \boldsymbol{\gamma}_3 = \boldsymbol{\alpha}_3$. We consider the cases when 20% (first and second panels), 50% (third and fourth panels) or 80% (fifth and sixth panels) of the α_{3j}'s are non -zero; and when 50% (first, third and fifth panels) or 100% (second, fourth and sixth panels) of α_{3j}'s are positive. $\boldsymbol{\alpha}_3$ controls the association between $\boldsymbol{G}$ and Y (main rare variants effects on Y). $\boldsymbol{\gamma}_3$ controls the association between $\boldsymbol{G}$ and E (main rare variants effects on E). $\boldsymbol{\beta}$ reflects the rare variants by environment interactions. For 71% of the 10^5 simulations, iSKAT with null model estimated without penalty ($\lambda = 0$) did not converge. Here, we only compare the two when iSKAT with null model estimated without penalty ($\lambda = 0$) converged. Power from both approaches are similar for the remaining 29% of simulations.

20% non-zero α_{3j}'s, 50% of α_{3j}'s are positive, $\boldsymbol{\beta} = \boldsymbol{\alpha}_3$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	7.02e-02	7.03e-02
1e-03	2.56e-02	2.59e-02

20% non-zero α_{3j}'s, 100% of α_{3j}'s are positive, $\boldsymbol{\beta} = \boldsymbol{\alpha}_3$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	7.96e-02	8.00e-02
1e-03	2.91e-02	2.94e-02

50% non-zero α_{3j}'s, 50% of α_{3j}'s are positive, $\boldsymbol{\beta} = \boldsymbol{\alpha}_3$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	1.84e-01	1.85e-01
1e-03	8.80e-02	8.85e-02

50% non-zero α_{3j}'s, 100% of α_{3j}'s are positive, $\boldsymbol{\beta} = \boldsymbol{\alpha}_3$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	2.62e-01	2.65e-01
1e-03	1.44e-01	1.45e-01

80% non-zero α_{3j}'s, 50% of α_{3j}'s are positive, $\boldsymbol{\beta} = \boldsymbol{\alpha}_3$

α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	3.03e-01	3.04e-01
1e-03	1.77e-01	1.78e-01

80% non-zero α_{3j}'s, 100% of α_{3j}'s are positive, $\boldsymbol{\beta} = \boldsymbol{\alpha}_3$

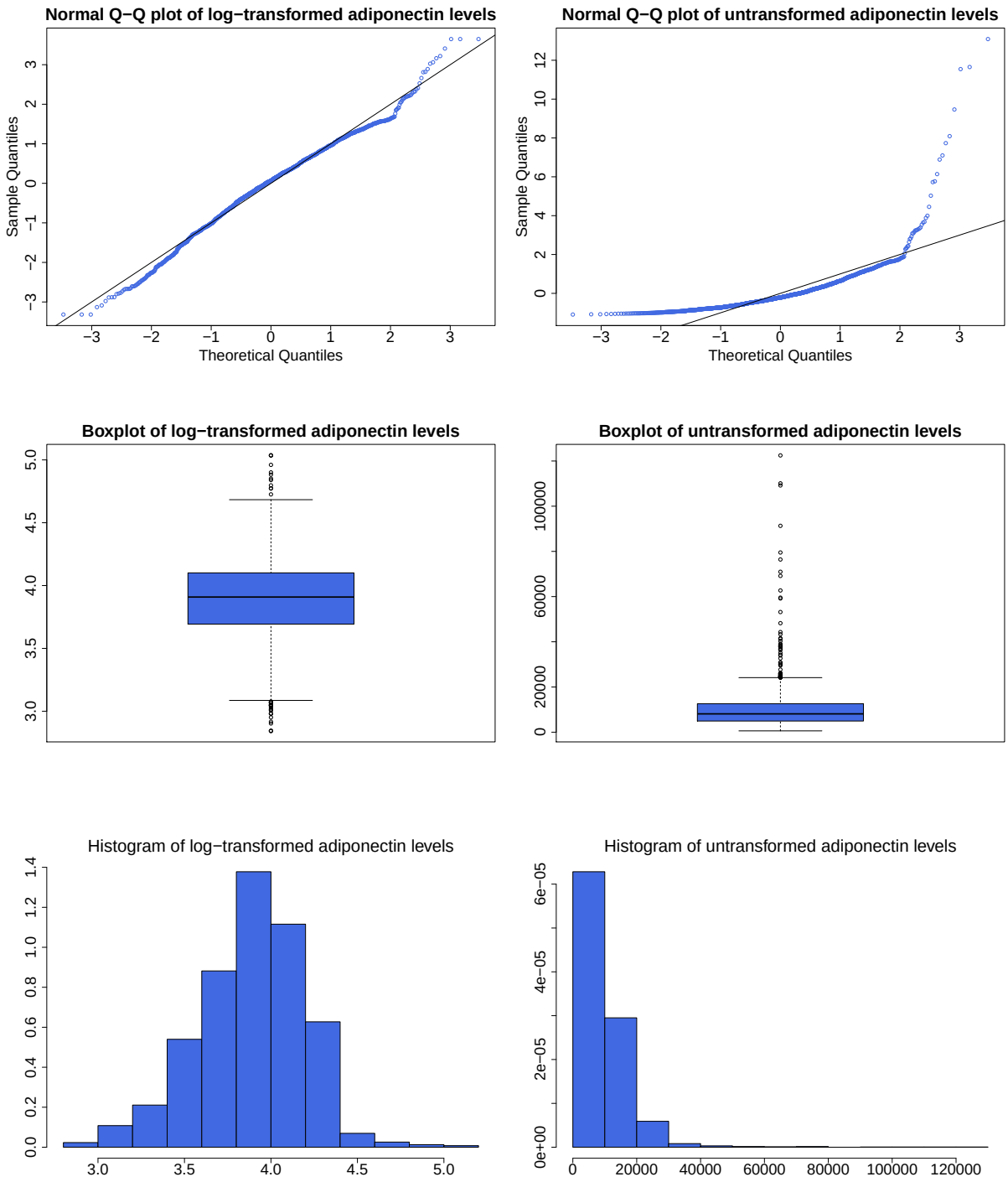
α-level	iSKAT ($\lambda = 0$)	iSKAT ($\lambda > 0$)
1e-02	4.71e-01	4.74e-01
1e-03	3.33e-01	3.35e-01

Web Table 6: Minor allele frequencies (MAF) and missing rates of each of the 11 rare variants within the adiponectin gene in the CoLaus resequencing dataset.

Variant	MAF	Missing Rates
chr3:188053510	2.15e-02	0.77%
chr3:188053533	2.58e-04	0.31%
chr3:188053663	2.58e-04	0.21%
chr3:188053705	5.16e-04	0.31%
chr3:188053732	2.57e-04	0.10%
chr3:188054673	2.61e-04	1.39%
chr3:188054720	5.95e-03	0.62%
chr3:188054724	2.58e-04	0.51%
chr3:188054783	2.57e-02	0.051%
chr3:188055047	2.58e-04	0.15%
chr3:188055252	3.68e-03	2.06%

Web Table 7: p-values from applying SKAT-O to the CoLaus resequencing dataset to assess rare variants main effects of the adiponectin gene on adiponectin levels. The SKAT-O model is a special case of the iSKAT model and corresponds to setting the ridge parameter $\lambda = \infty$ (i.e. $\boldsymbol{\alpha}_3 = \mathbf{0}$) and $\boldsymbol{S} = \boldsymbol{G}$ in the iSKAT model (main manuscript Equation (1)). p-values p_ρ (second column) correspond to each Q_ρ for each fixed $\rho = 0, 0.1, \cdots, 0.9, 1$ (first column). Each of the p_ρ corresponds to the p-value obtained if ρ was fixed *a priori*. The minimum of these 11 p-values gives the SKAT-O test statistic. SKAT-O, which accounts for multiple testing from using the minimum of these 11 p-values, gave a p-value of $1.8e-14$, confirming the strong association between rare variants within the adiponectin gene and adiponectin levels.

ρ	p_ρ
0	1.7e-15
0.1	4.5e-15
0.2	1.8e-13
0.3	1.3e-11
0.4	4.5e-10
0.5	6.8e-09
0.6	5.3e-08
0.7	2.7e-07
0.8	9.7e-07
0.9	2.8e-06
1	6.7e-06



Web Figure 31: Normal quantile-quantile plot (top panel), boxplot (middle panel) and histogram of transformed adiponectin levels (left panel) and untransformed adiponectin levels (right panel) respectively. For the left panel, a log10 transformation was applied to the plasma adiponectin levels and extreme values of adiponectin levels (values exceeding lower 0.1% or upper 99.9% percentile) were set to the boundary value (value at 0.1% or 99.9%), to improve normality and lessen the impact of outliers. Plots for raw values of adiponectin levels are shown on the right panel for comparison.

Web Table 8: p-values from applying iSKAT to the CoLaus resequencing dataset to assess rare variants by environment interactions. p-values p_ρ (second column) corresponding to each Q_ρ for each fixed $\rho = 0, 0.1, \cdots, 0.9, 1$ (first column) were obtained as described in Web Appendix 2.2. Each of the p_ρ corresponds to the p-value obtained if ρ was fixed *a priori*. The minimum of these 11 p-values gives the iSKAT test statistic (main manuscript Equation (8)). iSKAT, which accounts for multiple testing from using the minimum of these 11 p-values, gave a p-value of 0.037.

ρ	p_ρ
0	0.23
0.1	0.14
0.2	0.085
0.3	0.056
0.4	0.042
0.5	0.034
0.6	0.029
0.7	0.026
0.8	0.024
0.9	0.023
1	0.022

Web Table 9: Estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$ from weighted ridge regression (Web Appendix 2.1.). iSKAT uses these estimates from weighted ridge regression in the testing procedure. Estimated rare variants main effects from unpenalized regression are given in Web Table 10. Both sets of estimates are similar.

Variant	$\hat{\alpha}_{3j}$
chr3:188053510	-2.18e-01
chr3:188053533	-3.94e-01
chr3:188053663	-2.16e-02
chr3:188053705	-4.73e-01
chr3:188053732	-2.57e-01
chr3:188054673	1.22e-01
chr3:188054720	-1.51e-01
chr3:188054724	-8.37e-01
chr3:188054783	9.44e-02
chr3:188055047	-1.31e-02
chr3:188055252	-1.33e-01

Web Table 10: Estimated rare variants main effects $\hat{\boldsymbol{\alpha}}_3$, standard errors (SE) and p-values (P) from unpenalized linear regression. Unpenalized linear regression is equivalent to setting the ridge parameter ($\lambda = 0$) in the ridge regression framework (Web Appendix 2.1.). Estimated rare variants main effects from ridge regression are given in Web Table 9. Both sets of estimates are similar.

Variant	$\hat{\alpha}_{3j}$	SE	P
chr3:188053510	-2.18e-01	3.0e-02	8.6e-13
chr3:188053533	-3.97e-01	2.7e-01	1.4e-01
chr3:188053663	-2.18e-02	2.7e-01	9.4e-01
chr3:188053705	-4.76e-01	1.9e-01	1.4e-02
chr3:188053732	-2.59e-01	2.7e-01	3.4e-01
chr3:188054673	1.23e-01	2.7e-01	6.5e-01
chr3:188054720	-1.51e-01	5.7e-02	8.2e-03
chr3:188054724	-8.45e-01	2.7e-01	1.9e-03
chr3:188054783	9.45e-02	2.8e-02	6.1e-04
chr3:188055047	-1.32e-02	2.7e-01	9.6e-01
chr3:188055252	-1.33e-01	7.3e-02	6.7e-02

Web Table 11: Linkage disequilibrium (LD) measures ($D^{'}$ and R^2) between chr3:188053586 (MAF = 0.138 > 0.05) and the remainder 11 rare variants (MAF < 0.05) in the CoLaus resequencing dataset.

Variant	$D^{'}$	R^2
chr3:188053510	0.731	0.002
chr3:188053533	1	0
chr3:188053663	1	0
chr3:188053705	1	0.003
chr3:188053732	1	0
chr3:188054673	1	0
chr3:188054720	0.888	0.001
chr3:188054724	1	0.002
chr3:188054783	1	0.004
chr3:188055047	1	0
chr3:188055252	0.546	0

Web Table 12: Summary of association analysis results of the CoLaus resequencing dataset when missing genotypes were imputed by mean. Association analysis results when missing genotypes were imputed by homozygote of the major allele are given in main manuscript Table 2. Results from both sets of analysis are similar. The SKAT-O test (Lee *and others*, 2012) was used to test for the main rare variant effects on adiponectin levels (first row) and their effects on alcohol usage (second row). The iSKAT test was used to test for interaction effects between ADIPOQ gene and alcohol use on adiponectin levels (third-fifth rows).

Analysis	p-value
Main effects of rare variants of ADIPOQ gene on adiponectin levels	2.5e-14
Main effects of rare variants of ADIPOQ gene on alcohol usage	4.2e-02
Interaction effects of rare variants of ADIPOQ gene *alcohol on adiponectin levels	3.4e-02
Interaction effects of rare variants of ADIPOQ gene*alcohol on adiponectin levels, adjusting for effects of common variant chr3:188053586 and chr3:188053586*alcohol	6.1e-02
Interaction effects of ADIPOQ gene*alcohol (chr3:188053586*alcohol and rare variants*alcohol) on adiponectin levels	3.0e-02

Web Table 13: Summary of association analysis results of the CoLaus resequencing dataset, excluding singletons, when missing genotypes were imputed by homozygote of the major allele. Association analysis results without excluding singletons, when missing genotypes were imputed by homozygote of the major allele are given in main manuscript Table 2. Results from both sets of analysis are similar. The SKAT-O test (Lee *and others*, 2012) was used to test for the main rare variant effects on adiponectin levels (first row) and their effects on alcohol usage (second row). The iSKAT test was used to test for interaction effects between ADIPOQ gene and alcohol use on adiponectin levels (third-fifth rows).

Analysis	p-value
Main effects of rare variants of ADIPOQ gene on adiponectin levels	2.1e-14
Main effects of rare variants of ADIPOQ gene on alcohol usage	4.0e-02
Interaction effects of rare variants of ADIPOQ gene *alcohol on adiponectin levels	3.7e-02
Interaction effects of rare variants of ADIPOQ gene*alcohol on adiponectin levels, adjusting for effects of common variant chr3:188053586 and chr3:188053586*alcohol	6.4e-02
Interaction effects of ADIPOQ gene*alcohol (chr3:188053586*alcohol and rare variants*alcohol) on adiponectin levels	3.1e-02

References

Davies, R. (1980). The distribution of a linear combination of chi-square random variables. *Applied Statistics* **29**, 323–333.

Lee, S., Wu, M.C. and Lin, X. (2012). Optimal tests for rare variant effects in sequencing association studies. *Biostatistics* **31**, 762–775.

Lin, X., Lee, S., Christiani, D.C. and Lin, X. (2013). Test for interactions between a gene/snp-set and environment/treatment in generalized linear models. *Biostatistics* **14**, 667–681.

Liu, H., Tang, Y. and Zhang, H.H. (2009). A new chi-square approximation to the distribution of non-negative definite quadratic forms in non-central normal variables. *Computational Statistics & Data Analysis* **53**(4), 853–856.

O'Sullivan, F., Yandell, BS and Raynor Jr, WJ. (1986). Automatic smoothing of regression functions in generalized linear models. *Journal of the American Statistical Association* **81**(393), 96–103.

Schaffner, S.F., Foo, C., Gabriel, S., Reich, D., Daly, M.J. and Altshuler, D. (2005). Calibrating a coalescent simulation of human genome sequence variation. *Genome Research* **15**(11), 1576–1583.