Ensuring valid inference for hazard ratios after variable selection: Supplementary Materials

Kelly Van Lancker

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, U.S.A.

Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent, Belgium

email: kelly.vanlancker@ugent.be

and

Oliver Dukes

Department of Statistics and Data Science, The Wharton School, University of Pennsylvania, Philadelphia, U.S.A.

Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent, Belgium

email: oliver.dukes@ugent.be

and

Stijn Vansteelandt

Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent, Belgium

Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, U.K.

email: stijn.vansteelandt@ugent.be

Appendix A: Proof of Theorem 3.3

In order to prove Theorem 3.3, we will focus on the following score function

$$\hat{U}_i(\alpha,\beta,\gamma) = \int_0^\tau \left[A_i - \bar{A}_n(t,\alpha,\beta) - \gamma'\left\{L_i - \bar{L}_n(t,\alpha,\beta)\right\}\right]\{dN_i(t) - R_i(t)\hat{\lambda}_0(t,\alpha,\beta)e^{\alpha A_i+\beta' L_i}dt\},$$

where $\bar{A}_n(t,\alpha,\beta) = \frac{\mathbb{E}_n\left\{AR(t)e^{\alpha A+\beta' L}\right\}}{\mathbb{E}_n\left\{R(t)e^{\alpha A+\beta' L}\right\}}$, $\bar{L}_n(t,\alpha,\beta) = \frac{\mathbb{E}_n\left\{LR(t)e^{\alpha A+\beta' L}\right\}}{\mathbb{E}_n\left\{R(t)e^{\alpha A+\beta' L}\right\}}$ and $\hat{\lambda}_0(t,\alpha,\beta) = \frac{\mathbb{E}_n\{R(t)dN(t)\}}{\mathbb{E}_n\left\{R(t)e^{\alpha A+\beta' L}\right\}}$. We denote the population versions of the above-defined score function by

$$U_i^*(\alpha,\beta,\gamma) = \int_0^\tau \left[A_i - \bar{A}^*(t,\alpha,\beta) - \gamma'\left\{L_i - \bar{L}^*(t,\alpha,\beta)\right\}\right]\{dN_i(t) - R_i(t)\lambda_{0n}(t,\alpha,\beta)e^{\alpha A_i+\beta' L_i}dt\},$$

where $\bar{A}^*(t,\alpha,\beta) = \frac{\mathbb{E}_{P_n}\left\{AR(t)e^{\alpha A+\beta' L}\right\}}{\mathbb{E}_{P_n}\left\{R(t)e^{\alpha A+\beta' L}\right\}}$, $\bar{L}^*(t,\alpha,\beta) = \frac{\mathbb{E}_{P_n}\left\{LR(t)e^{\alpha A+\beta' L}\right\}}{\mathbb{E}_{P_n}\left\{R(t)e^{\alpha A+\beta' L}\right\}}$ and $\lambda_{0n}(t,\alpha,\beta) = \frac{\mathbb{E}_{P_n}\{R(t)dN(t)\}}{\mathbb{E}_{P_n}\left\{R(t)e^{\alpha A+\beta' L}\right\}}$. Moreover, we will rely on the primitive conditions required for uniform consistency of the Lasso estimators (see e.g., Assumptions 2 and 5 in Fang et al. (2017)) and the following assumptions, where Assumption 1 corresponds with Assumption 1 in Fang et al. (2017) and Assumption 2 with Assumption 3 and 4 in the same paper.

Assumption 1 (Bounded covariates):

$$\max_{i\leqslant n} ||L_i||_\infty \leqslant C_L < +\infty,$$

with $||Z||_\infty = \max_{j\leqslant p}|Z_j|$ for any vector Z.

Assumption 2 (Bounded population predictions): It holds that $s_\gamma \asymp s_\beta \asymp s_\eta$ (i.e., s_γ, s_β and s_η have the same order of magnitude) and

(1) $\max_{i\leqslant n}|\beta_n' L_i| = O_{P_n}(1)$

(2) $\max_{i\leqslant n}|\gamma_n' L_i| = O_{P_n}(1)$

(3) $\max_{i\leqslant n}|\eta_{2n}' L_i| = O_{P_n}(1)$

Assumption 3:

$$\min_{0\leqslant t\leqslant\tau}\left|\frac{1}{n}\sum_{i=1}^n R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\right| \geqslant \delta > 0$$

Assumption 3 essentially states that the average of the population hazard ratios $e^{\alpha_n A_i + \beta_n' L_i}$ for people at risk is bounded away from zero at each time point t ($\in [0, \tau]$). This means that τ should be kept sufficiently small so that there is a remaining risk to fail beyond τ.

For what follows, let $X = (A, L')'$, $\theta = (\alpha, \beta')'$ and correspondingly $\theta_n = (\alpha_n, \beta_n')'$, $\hat{\theta} = (\hat{\alpha}, \hat{\beta}')'$ and $\check{\theta} = (\check{\alpha}, \check{\beta}')'$. We will make use of the following rate conditions

Assumption 4 (Rates on error of estimated coefficients):

(1) $||\theta_n - \hat{\theta}||_1 = O_{P_n}\left(s_\theta \sqrt{\log(p \vee n)/n}\right)$
(2) $||\theta_n - \hat{\theta}||_2 = O_{P_n}\left(\sqrt{s_\theta \log(p \vee n)/n}\right)$
(3) $||\gamma_n - \hat{\gamma}||_1 = O_{P_n}\left(s_\gamma \sqrt{\log(p \vee n)/n}\right)$
(4) $||\gamma_n - \hat{\gamma}||_2 = O_{P_n}\left(\sqrt{s_\gamma \log(p \vee n)/n}\right)$
(5) $||\eta_{2n} - \hat{\eta}_2||_1 = O_{P_n}\left(s_\eta \sqrt{\log(p \vee n)/n}\right)$
(6) $||\eta_{2n} - \hat{\eta}_2||_2 = O_{P_n}\left(\sqrt{s_\eta \log(p \vee n)/n}\right)$

Here, $\hat{\gamma}$ and $\hat{\eta}_2$ are the post-Lasso estimates of γ and η_2 respectively, obtained by refitting the model in respectively Step 3 and Step 2 on all covariates selected by the Lasso step (in respectively Step 3 and 2).

Remark 1: Rates of convergence 1, 2, 5 and 6 are shown to hold for Lasso estimates in Huang et al. (2013) under Assumption 1 and a compatibility factor condition. The validity of the latter assumption, which bounds the minimal eigenvalues of the Hessian matrix

$$-\frac{1}{n}\sum_{i=1}^{n}\int_0^\tau \left[X_i - \frac{\mathbb{E}_n\left\{XR(t)e^{\alpha_n A + \beta_n' L}\right\}}{\mathbb{E}_n\left\{R(t)e^{\alpha_n A + \beta_n' L}\right\}}\right] \mathrm{d}N_i(t)$$

from below, has been verified in Theorem 4.1 of Huang et al. (2013). For post-Lasso estimates we refer the reader to Belloni and Chernozhukov (2013). Rates 3 and 4 in Assumption 4 follow from the results of Belloni et al. (2016) on weighted l_1-penalized regression and Fang et al. (2017). We assume the primitive conditions required for uniform consistency of the Lasso estimators to hold (see e.g., Belloni et al., 2012).

As support($\hat{\gamma}$) $\subseteq B$, support($\hat{\beta}$) $\subseteq B$ and support($\hat{\eta}$) $\subseteq B$, one can show that $||\theta_n - \check{\theta}||_1 = O_{P_n}\left(s_\theta\sqrt{\log(p\vee n)/n}\right)$, $||\theta_n - \check{\theta}||_2 = O_{P_n}\left(\sqrt{s_\theta \log(p\vee n)/n}\right)$, $||\gamma_n - \check{\gamma}||_1 = O_{P_n}\left(s_\gamma\sqrt{\log(p\vee n)/n}\right)$, $||\gamma_n - \check{\gamma}||_2 = O_{P_n}\left(\sqrt{s_\gamma \log(p\vee n)/n}\right)$, $||\eta_{2n} - \check{\eta}_2||_1 = O_{P_n}\left(s_\eta\sqrt{\log(p\vee n)/n}\right)$ and $||\eta_{2n} - \check{\eta}_2||_2 = O_{P_n}\left(\sqrt{s_\eta \log(p\vee n)/n}\right)$ (see Belloni and Chernozhukov, 2013).

To obtain the rates in Assumption 4 we need that

$$\lambda_\theta = O\left(\sqrt{\frac{\log(p\vee n)}{n}}\right) \tag{A.1}$$

$$\lambda_\gamma = O\left(\sqrt{\frac{\log(p\vee n)}{n}}\right) \tag{A.2}$$

$$\lambda_\eta = O\left(\sqrt{\frac{\log(p\vee n)}{n}}\right), \tag{A.3}$$

where λ_θ was denoted before as $\lambda_{\alpha,\beta}$ (see Equation (3.7) in the main paper) and $\lambda_\eta > 0$ is a Lasso penalty parameter for the censoring model in Step 2 of the selection procedure. These are standard assumptions on the order of the penalty level (Bühlmann and van de Geer, 2011).

Assumption 5 (Concentration of the gradients): There exist positive constants C_γ and C_θ, such that,

(1) $\left|\left|\mathbb{E}_n\left\{\frac{\partial \hat{U}_i(\alpha_n,\beta_n,\gamma_n)}{\partial\gamma}\right\}\right|\right|_\infty \leqslant C_\gamma\sqrt{\frac{\log(p\vee n)}{n}}$

(2) $\left|\left|\mathbb{E}_n\left\{\frac{\partial \hat{U}_i(\alpha_n,\beta_n,\gamma_n)}{\partial\beta}\right\}\right|\right|_\infty \leqslant C_\beta\sqrt{\frac{\log(p\vee n)}{n}}$

In Fang et al. (2017) it is proven that, under Assumptions 1, 2 and primitive conditions required for uniform consistency of the Lasso estimators, Assumption 5.1 holds with probability at least $1 - O(p^{-3})$ and Assumption 5.2 with probability at least $1 - O(p^{-1})$ (for $p > n$).

Assumption 6:

$$\begin{aligned}&\left|\left|\frac{1}{n}\sum_{i=1}^n \int_0^\tau \left[A_i - \bar{A}_n(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\}\right]\left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\}\right.\right.\\ &\times\left.\left.\left\{dN_i(t) - R_i(t)e^{\alpha_n A_i + \beta_n' L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)dt\right\}\right|\right|_\infty = O_{P_n}\left(\sqrt{\frac{\log(p\vee n)}{n}}\right).\end{aligned}$$

We conjecture Assumption 6 to hold, under a correctly specified Cox model for survival time T, based on similar results from an extension of the moderate deviation theory for self-normalized sums (de la Peña et al., 2009). Consequently, if Assumption 6 holds then Assumption 5.2 is equivalent to assuming that

$$\begin{aligned}&\left\|\mathbb{E}_n\left(\int_0^\tau \left[A_i - \bar{A}_n(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\}\right]\right.\right.\\&\left.\left.\times\left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\}dN_i(t)\right)\right\|_\infty = O_{P_n}\left(\sqrt{\frac{\log(p\vee n)}{n}}\right).\end{aligned} \tag{A.4}$$

Note that if we plug the Lasso estimates from Step 1 and Step 3 in Assumption 5.1 and Assumption (A.4), connecting to the maximum penalized log-likelihood, by the Karush-Kuhn-Tucker (KKT) conditions these gradients will be upper bounded by respectively λ_θ and λ_γ (Bühlmann and van de Geer, 2011).

Assumption 7:

(1) $\sup_{t\in[0,\tau]}\left\|\mathbb{E}_{P_n}\left\{R(t)Xe^{\alpha_n A+\beta_n' L}\right\} - \mathbb{E}_n\left\{R(t)Xe^{\alpha_n A+\beta_n' L}\right\}\right\|_\infty = O_{P_n}\left(\sqrt{\frac{\log(p\vee n)}{n}}\right)$

(2) $\sup_{t\in[0,\tau]}\left|\mathbb{E}_{P_n}\left\{R(t)e^{\alpha_n A+\beta_n' L}\right\} - \mathbb{E}_n\left\{R(t)e^{\alpha_n A+\beta_n' L}\right\}\right| = O_{P_n}\left(\sqrt{\frac{\log(p\vee n)}{n}}\right)$.

In Fang et al. (2017) it is shown that this assumption holds under Assumptions 1, 2, 3 and primitive conditions required for uniform consistency of the Lasso estimators (for $p > n$).

Assumption 8: For all $t \in [0,\tau]$

(1) $\mathbb{E}_n\left[(\gamma_n - \hat{\gamma})'\left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\}\right]^2 = O_{P_n}\left(\frac{s_\gamma\log(p\vee n)}{n}\right)$,

(2) $\mathbb{E}_n\left[(\beta_n - \hat{\beta})' L_i\right]^2 = O_{P_n}\left(\frac{s_\theta\log(p\vee n)}{n}\right)$.

Although these assumptions are not proven to hold, we judge them to be plausible. For example, in Lemler (2016) and Bertin et al. (2011) it is shown that Assumption 8 is reasonable in high-dimensional settings (for $p > n$) where $\beta' L$ and the baseline hazard are jointly estimated. As we do not need the baseline hazard in the testing procedure described in Section 3 of the main paper, this condition is even more likely satisfied in our case.

Finally, we will assume the standard regularity conditions to interchange the order of summation and integration.

Proof. [Proof of Theorem 3.3] This proof consists of three parts. In the first part we prove the asymptotic normality of $\sqrt{n}(\tilde{\alpha} - \alpha_n)$, i.e., $\sqrt{n}(\tilde{\alpha} - \alpha_n) \xrightarrow{d} N(0, \Sigma_n^2)$, where $\tilde{\alpha}$ is the solution to $\mathbb{E}_n \left\{ \hat{U}_i \left(\alpha, \hat{\beta}, \hat{\gamma} \right) \right\} = 0$. Note that $\tilde{\alpha}$ is closely related to the one-step estimator proposed in Fang et al. (2017), the main difference being that post-Lasso (rather than Lasso) estimates for β and γ are used. The proof of the first part closely follows the proof of their Theorem 3.3. In the second part we show the consistency of the variance estimator, evaluated at $\tilde{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$. In the final part, we then relate the results from the two previous parts to the proposed triple selection approach.

Part 1: Normality

We first consider the sample mean of $\hat{U}_i(\tilde{\alpha}, \hat{\beta}, \hat{\gamma})$:

$$\mathbb{E}_n\{\hat{U}_i(\tilde{\alpha}, \hat{\beta}, \hat{\gamma})\} = \mathbb{E}_n\{U_i^*(\alpha_n, \beta_n, \gamma_n) + \hat{U}_i(\tilde{\alpha}, \hat{\beta}, \hat{\gamma}) - U_i^*(\alpha_n, \beta_n, \gamma_n)\}.$$

After some algebra, we have

$$\sqrt{n}\mathbb{E}_n\{\hat{U}_i(\tilde{\alpha}, \hat{\beta}, \hat{\gamma})\} = \sqrt{n}\mathbb{E}_n\{U_i^*(\alpha_n, \beta_n, \gamma_n)\} + R_1 + R_2 + R_3 + R_4 + R_5, \tag{A.5}$$

where

$$\begin{aligned}
R_1 &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^{\tau} \left(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right. \\
&\quad \left. \times \left\{ \lambda_{0n}(t, \alpha_n, \beta_n) - \hat{\lambda}_0(t, \alpha_n, \beta_n) \right\} \right) dt, \\
R_2 &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^{\tau} \left(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) \right. \\
&\quad \left. \times \left\{ \hat{\lambda}_0(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} - \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta}) e^{\tilde{\alpha} A_i + \hat{\beta}' L_i} \right\} \right) dt, \\
R_3 &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^{\tau} R_i(t) \left\{ dN_i(t) - \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta}) e^{\tilde{\alpha} A_i + \hat{\beta}' L_i} dt \right\} \left\{ \bar{A}_n(t, \tilde{\alpha}, \hat{\beta}) - \bar{A}^*(t, \alpha_n, \beta_n) \right\}, \\
R_4 &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^{\tau} R_i(t) \left\{ dN_i(t) - \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta}) e^{\tilde{\alpha} A_i + \hat{\beta}' L_i} dt \right\} \gamma_n' \left\{ \bar{L}^*(t, \alpha_n, \beta_n) - \bar{L}_n(t, \tilde{\alpha}, \hat{\beta}) \right\},
\end{aligned}$$

$$R_5 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{ L_i - \bar{L}_n(t, \tilde{\alpha}, \hat{\beta}) \right\} \left\{ dN_i(t) - \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta}) e^{\tilde{\alpha} A_i + \hat{\beta}' L_i} dt \right\}.$$

First, consider R_1,

$$\begin{aligned}
&\left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right. \\
&\left. \times \left\{ \lambda_{0n}(t, \alpha_n, \beta_n) - \hat{\lambda}_0(t, \alpha_n, \beta_n) \right\} dt \right| \\
&= \sqrt{n} \left| \int_0^\tau \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right. \right. \\
&\left. \left. \times \left[\lambda_{0n}(t, \alpha_n, \beta_n) - \frac{\mathbb{E}_n \left\{ R(t) dN(t) \right\}}{\mathbb{E}_n \left\{ R(t) e^{\alpha_n A + \beta_n' L} \right\}} \right] dt \right) \right| \\
&= \sqrt{n} \left| \int_0^\tau \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right. \right. \\
&\left. \left. \times \frac{\mathbb{E}_n \left[R(t) \{ \lambda_{0n}(t, \alpha_n, \beta_n) e^{\alpha_n A + \beta_n' L} dt - dN(t) \} \right]}{\mathbb{E}_n \left\{ R(t) e^{\alpha_n A + \beta_n' L} \right\}} \right) \right| \\
&= \sqrt{n} \left| \int_0^\tau \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right. \right. \\
&\left. \left. \times \frac{\mathbb{E}_n \left\{ dM(t) \right\}}{\mathbb{E}_n \left\{ R(t) e^{\alpha_n A + \beta_n' L} \right\}} \right) \right| \\
&\leqslant \sqrt{n} \int_0^\tau \left| \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right) \right| \\
&\times \left| \frac{\mathbb{E}_n \left\{ dM(t) \right\}}{\mathbb{E}_n \left\{ R(t) e^{\alpha_n A + \beta_n' L} \right\}} \right| \\
&\leqslant \sqrt{n} \sup_{t \in [0, \tau]} \left| \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right) \right| \\
&\times \int_0^\tau \left| \frac{\mathbb{E}_n \left\{ dM(t) \right\}}{\mathbb{E}_n \left\{ R(t) e^{\alpha_n A + \beta_n' L} \right\}} \right| \\
&\leqslant \sqrt{n} \delta^{-1} \sup_{t \in [0, \tau]} \left| \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right) \right| \\
&\times \int_0^\tau \left| \mathbb{E}_n \left\{ dM(t) \right\} \right|
\end{aligned}$$

where the last inequality follows from Assumption 3. From the proof of Lemma 3 in Fang et al. (2017) it follows that $\int_0^\tau \left| \frac{1}{n} \sum_{j=1}^n dM_j(t) \right| = O_{P_n}(n^{-1/2})$ under a correctly specified Cox model for survival time T. We conclude that, for $C_n = O_{P_n}(1)$,

$$|R_1| \leqslant C_n \delta^{-1} \sup_{t\in[0,\tau]} \left| \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right) \right|.$$

Under Assumption 2, it follows from the uniform weak law of large numbers that

$$\begin{aligned}
\forall t \in [0,\tau] : & \left| \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right) \right| \\
= & \left| \mathbb{E}_n \left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right) \right. \\
& \left. - \mathbb{E}_{P_n} \left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right) \right| \\
= & O_{P_n}(n^{-1/2}),
\end{aligned}$$

since $\mathbb{E}_{P_n}\left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right) = 0$ by the definitions of $\bar{A}^*(t,\alpha_n,\beta_n)$ and $\bar{L}^*(t,\alpha_n,\beta_n)$. By the definition of a supremum, it then follows that $|R_1| = o_{P_n}(1)$.

Now, let $\bar{X}^*(t,\alpha,\beta) = \frac{\mathbb{E}_{P_n}\left\{ R(t) X e^{\alpha A + \beta' L} \right\}}{\mathbb{E}_{P_n}\left\{ R(t) e^{\alpha A + \beta' L} \right\}}$ and $\bar{X}_n(t,\alpha,\beta) = \frac{\mathbb{E}_n\left\{ R(t) X e^{\alpha A + \beta' L} \right\}}{\mathbb{E}_n\left\{ R(t) e^{\alpha A + \beta' L} \right\}}$. Then, following a Taylor expansion for R_2 with respect to $\tilde{\theta} = (\tilde{\alpha}, \hat{\beta}')'$ yields

$$\begin{aligned}
& \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau \left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) \right. \\
& \left. \times \{ \hat{\lambda}_0(t,\alpha_n,\beta_n) e^{\alpha_n A_i + \beta_n' L_i} - \hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta}) e^{\tilde{\alpha} A_i + \hat{\beta}' L_i} \} dt \right) \\
= & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau \left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t,\alpha_n,\beta_n) \right. \\
& \left. \times \left\{ X_i - \bar{X}_n(t,\alpha_n,\beta_n) \right\} dt \right) (\theta_n - \tilde{\theta}) + O_{P_n}(\sqrt{n} ||\theta_n - \tilde{\theta}||_2^2) \\
= & \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau \left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t,\alpha_n,\beta_n) \right. \\
& \left. \times \left\{ A_i - \bar{A}_n(t,\alpha_n,\beta_n) \right\} dt \right) (\alpha_n - \tilde{\alpha}) \\
& + \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau \left(\left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t,\alpha_n,\beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t,\alpha_n,\beta_n) \right. \\
& \left. \times \left\{ L_i - \bar{L}_n(t,\alpha_n,\beta_n) \right\} dt \right) (\beta_n - \hat{\beta}) \\
& + O_{P_n}(\sqrt{n} ||\theta_n - \tilde{\theta}||_2^2),
\end{aligned}$$

where

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^\tau \Big(\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)$$

$$\times\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}dt\Big)(\beta_n-\hat{\beta})$$

$$=\sqrt{n}\mathbb{E}_n\left[\int_0^\tau \Big(\left[A_i-\bar{A}_n(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right.$$

$$\left.\times\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}dt\Big)\right](\beta_n-\hat{\beta}) \tag{A.6}$$

$$+\sqrt{n}\mathbb{E}_n\left[\int_0^\tau \left\{\bar{A}_n(t,\alpha_n,\beta_n)-\bar{A}^*(t,\alpha_n,\beta_n)\right\}R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right.$$

$$\left.\times\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}dt\right](\beta_n-\hat{\beta}) \tag{A.7}$$

$$+\sqrt{n}\gamma_n'\mathbb{E}_n\left[\int_0^\tau \left\{\bar{L}^*(t,\alpha_n,\beta_n)-\bar{L}_n(t,\alpha_n,\beta_n)\right\}R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right.$$

$$\left.\times\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}dt\right](\beta_n-\hat{\beta}). \tag{A.8}$$

Term (A.6) yields

$$\Bigg|\sqrt{n}\mathbb{E}_n\left[\int_0^\tau \Big(\left[A_i-\bar{A}_n(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right.$$

$$\left.\times\ \left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}dt\Big)\right](\beta_n-\hat{\beta})\Bigg|$$

$$\leqslant\sqrt{n}\left|\left|\mathbb{E}_n\left[\int_0^\tau \Big(\left[A_i-\bar{A}_n(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right.\right.\right.$$

$$\left.\left.\left.\times\ \left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}dt\Big)\right]\right|\right|_\infty \left|\left|\beta_n-\hat{\beta}\right|\right|_1$$

$$=\sqrt{n}\left|\left|\mathbb{E}_n\left\{\frac{\partial\hat{U}_i(\alpha_n,\beta_n,\gamma_n)}{\partial\beta}\right\}\right|\right|_\infty\cdot\left|\left|\beta_n-\hat{\beta}\right|\right|_1.$$

By Assumption 5.2, Assumption 4.1 and the ultra-sparsity condition in Theorem 3.3 we then have that

$$\Bigg|\sqrt{n}\mathbb{E}_n\left[\int_0^\tau \Big(\left[A_i-\bar{A}_n(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right.$$

$$\left.\times\ \left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}dt\Big)\right](\beta_n-\hat{\beta})\Bigg|=o_{P_n}(1).$$

Next, for term (A.7) we have

$$\sqrt{n}\mathbb{E}_n\left[\int_0^\tau \left\{\bar{A}_n(t,\alpha_n,\beta_n)-\bar{A}^*(t,\alpha_n,\beta_n)\right\}R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right.$$

$$\times\left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\}dt\Big](\beta_n-\hat{\beta})$$

$$=\sqrt{n}\int_0^\tau\Bigg\{\left\{\bar{A}_n(t,\alpha_n,\beta_n)-\bar{A}^*(t,\alpha_n,\beta_n)\right\}$$

$$\times\,\mathbb{E}_n\left(R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\left[L_i-\frac{\mathbb{E}_n\{R(t)Le^{\alpha_nA+\beta_n'L}\}}{\mathbb{E}_n\{R(t)e^{\alpha_nA+\beta_n'L}\}}\right]\right)dt\Bigg\}(\beta_n-\hat{\beta})$$

$$=0.$$

Similarly, term (A.8) equals zero. Thus, by Assumption 4.2 and the consistency of $\tilde{\alpha}$,

$$\begin{aligned}R_2=&\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau\Big(\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)\\&\times e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\left\{A_i-\bar{A}_n(t,\alpha_n,\beta_n)\right\}dt\Big)(\alpha_n-\tilde{\alpha})+o_{P_n}(1).\end{aligned}\tag{A.9}$$

Next, for R_3, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau R_i(t)\{dN_i(t)-\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}dt\}\left\{\bar{A}_n(t,\tilde{\alpha},\hat{\beta})-\bar{A}^*(t,\alpha_n,\beta_n)\right\}$$

$$=\sqrt{n}\mathbb{E}_n\left(\int_0^\tau R_i(t)\left[dN_i(t)-e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\frac{\mathbb{E}_n\left\{dN(t)R(t)\right\}}{\mathbb{E}_n\left\{R(t)e^{\tilde{\alpha}A+\hat{\beta}'L}\right\}}\right]\left\{\bar{A}_n(t,\tilde{\alpha},\hat{\beta})-\bar{A}^*(t,\alpha_n,\beta_n)\right\}\right)$$

$$=0.$$

Similarly, for R_4, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau R_i(t)\{dN_i(t)-\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}dt\}\gamma_n'\left\{\bar{L}^*(t,\alpha_n,\beta_n)-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}$$

$$=\sqrt{n}\mathbb{E}\left(\int_0^\tau R_i(t)\left[dN_i(t)-e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\frac{\mathbb{E}_n\left\{dN(t)R(t)\right\}}{\mathbb{E}_n\left\{R(t)e^{\tilde{\alpha}A+\hat{\beta}'L}\right\}}\right]\gamma_n'\left\{\bar{L}^*(t,\alpha_n,\beta_n)-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}\right)$$

$$=0.$$

Next, for R_5, note that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau R_i(t)(\gamma_n-\hat{\gamma})'\left\{L_i-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}\{dN_i(t)-\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}dt\}$$

$$=\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau R_i(t)(\gamma_n-\hat{\gamma})'\left\{L_i-\bar{L}_n(t,\alpha_n,\beta_n)\right\}\{dN_i(t)-\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}dt\},$$

as

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^\tau\Big(R_i(t)(\gamma_n-\hat{\gamma})'\left\{\bar{L}_n(t,\alpha_n,\beta_n)-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}$$

$$\times \left\{dN_i(t) - \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta})e^{\tilde{\alpha}A_i + \hat{\beta}' L_i} dt\right\}\Big) = 0$$

by a similar reasoning as for R_4.

Moreover,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\} \left\{dN_i(t) - \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta})e^{\tilde{\alpha}A_i + \hat{\beta}' L_i} dt\right\}$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\} \left\{dN_i(t) - \hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i} dt\right\} \quad \text{(A.10)}$$

$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^\tau \Big(R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}$$

$$\times \left\{\hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i} - \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta})e^{\tilde{\alpha}A_i + \hat{\beta}' L_i} dt\right\}\Big) . \quad \text{(A.11)}$$

Then, for term (A.10), by Hölder's inequality,

$$\left|\sqrt{n}\mathbb{E}_n \left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\} \left\{dN_i(t) - \hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i} dt\right\}\right]\right|$$

$$\leqslant \sqrt{n} \left|\left|\mathbb{E}_n \left[\int_0^\tau R_i(t) \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\} \left\{dN_i(t) - \hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i} dt\right\}\right]\right|\right|_\infty ||\gamma_n - \hat{\gamma}||_1$$

$$= \sqrt{n} \left|\left|\mathbb{E}_n \left\{\frac{\partial \hat{U}_i(\alpha_n, \beta_n, \gamma_n)}{\partial \gamma}\right\}\right|\right|_\infty ||\gamma_n - \hat{\gamma}||_1$$

$$\leqslant \sqrt{n} C_\gamma \sqrt{\frac{\log(p \vee n)}{n}} \, ||\gamma_n - \hat{\gamma}||_1 ,$$

using Assumption 5.1. Therefore, given Assumption 4.3 and the ultra-sparsity condition, this term converges in probability to zero.

A further Taylor expansion on term (A.11) yields

$$\sqrt{n}\mathbb{E}_n \left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}\right.$$

$$\left.\times \left\{\hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i} - \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta})e^{\tilde{\alpha}A_i + \hat{\beta}' L_i} dt\right\}\right]$$

$$= \sqrt{n}\mathbb{E}_n \left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\} \hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i}\right.$$

$$\left.\times \left\{A_i - \bar{A}_n(t, \alpha_n, \beta_n)\right\} (\alpha_n - \tilde{\alpha}) dt\right] \quad \text{(A.12)}$$

$$+ \sqrt{n}\mathbb{E}_n \left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\} \hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i}\right.$$

$$
\begin{aligned}
&\times \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}(\beta_n - \hat{\beta})dt\Big] \qquad \text{(A.13)} \\
&+ O_{P_n}(\sqrt{n}||\theta_n - \tilde{\theta}||_2^2).
\end{aligned}
$$

For Term (A.13) we have

$$
\begin{aligned}
&\sqrt{n}\left|\mathbb{E}_n\left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})'\left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}\hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i}\right.\right. \\
&\left.\left.\times \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}(\beta_n - \hat{\beta})dt\right]\right| \\
&\leqslant \sqrt{n}\tau \sup_{t\in[0,\tau]}\left|\mathbb{E}_n\left[R_i(t)(\gamma_n - \hat{\gamma})'\left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}\hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i}\right.\right. \\
&\left.\left.\times \left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}(\beta_n - \hat{\beta})dt\right]\right| \\
&\leqslant \sqrt{n}\tau \sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left[R_i(t)(\gamma_n - \hat{\gamma})'\left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}\hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i}\right]^2} \\
&\times \sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left[(\beta_n - \hat{\beta})'\left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}\right]^2} \\
&\leqslant \sqrt{n}\tau \sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left[R_i(t)(\gamma_n - \hat{\gamma})'\left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}\hat{\lambda}_0(t, \alpha_n, \beta_n)e^{\alpha_n A_i + \beta_n' L_i}\right]^2} \\
&\times \sqrt{2}\left[\sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left\{(\beta_n - \hat{\beta})'L_i\right\}^2} + \sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left\{(\beta_n - \hat{\beta})'\bar{L}_n(t, \alpha_n, \beta_n)\right\}^2}\right] \\
&\leqslant \sqrt{n}\tau C \sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left[(\gamma_n - \hat{\gamma})'\left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}\right]^2} \\
&\times \left[\sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left\{(\beta_n - \hat{\beta})'L_i\right\}^2} + \delta^{-1}\sup_{t\in[0,\tau]}\left|\mathbb{E}_n\left\{(\beta_n - \hat{\beta})'L_i R_i(t)e^{\alpha_n A_i + \beta_n' L_i}\right\}\right|\right]
\end{aligned}
$$

$$
\begin{aligned}
&\leqslant \sqrt{n}\tau C \sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left[(\gamma_n - \hat{\gamma})'\left\{L_i - \bar{L}_n(t, \alpha_n, \beta_n)\right\}\right]^2} \\
&\times \left[\sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left\{(\beta_n - \hat{\beta})'L_i\right\}^2} + \delta^{-1}C'\sup_{t\in[0,\tau]}\sqrt{\mathbb{E}_n\left\{(\beta_n - \hat{\beta})'L_i\right\}^2}\right],
\end{aligned}
$$

for positive constants C and C' by $(a+b)^2 \leqslant 2(a^2+b^2)$, Hölder's inequality, Assumption 2.1 and Assumption 3. Under Assumption 8, term (A.13) then converges in probability to zero.

It then follows by Assumption 4.2 and the consistency of $\tilde{\alpha}$ that

$$R_5 = \sqrt{n}\mathbb{E}_n \left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{ L_i - \bar{L}_n(t, \alpha_n, \beta_n) \right\} \hat{\lambda}_0(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} \right. \\ \left. \times \left\{ A_i - \bar{A}_n(t, \alpha_n, \beta_n) \right\} (\alpha_n - \tilde{\alpha}) dt \right] + o_{P_n}(1). \tag{A.14}$$

Next, it follows from the fact that $R_1 + R_3 + R_4 = o_{P_n}(1)$ and Equations (A.9) and (A.14), that Equation (A.5) reduces to

$$\sqrt{n}\mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}_n(t, \tilde{\alpha}, \hat{\beta}) - \hat{\gamma}' \left\{ L_i - \bar{L}_n(t, \tilde{\alpha}, \hat{\beta}) \right\} \right] \left\{ dN_i(t) - R_i(t) \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta}) e^{\tilde{\alpha} A_i + \hat{\beta}' L_i} dt \right\} \right)$$

$$= \sqrt{n}\mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] \right.$$

$$\left. \times \left\{ dN_i(t) - R_i(t) \lambda_{0n}(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} dt \right\} \right)$$

$$+ \sqrt{n}\mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t, \alpha_n, \beta_n) \right.$$

$$\left. \times \left\{ A_i - \bar{A}_n(t, \alpha_n, \beta_n) \right\} dt \right) (\alpha_n - \tilde{\alpha})$$

$$+ \sqrt{n}\mathbb{E}_n \left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{ L_i - \bar{L}_n(t, \alpha_n, \beta_n) \right\} \hat{\lambda}_0(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} \left\{ A_i - \bar{A}_n(t, \alpha_n, \beta_n) \right\} dt \right]$$

$$\times (\alpha_n - \tilde{\alpha})$$

$$+ o_{P_n}(1).$$

Since the estimator $\tilde{\alpha}$ can be defined as the solution to the estimating equations

$$\sum_{i=1}^n \int_0^\tau \left[A_i - \bar{A}_n(t, \alpha, \hat{\beta}) - \hat{\gamma}' \left\{ L_i - \bar{L}_n(t, \alpha, \hat{\beta}) \right\} \right] \left\{ dN_i(t) - R_i(t) \hat{\lambda}_0(t, \alpha, \hat{\beta}) e^{\alpha A_i + \hat{\beta}' L_i} dt \right\} = 0,$$

the left hand side of the previous equation becomes zero such that

$$0 = \sqrt{n}\mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] \right. \tag{A.15}$$

$$\left. \times \left\{ dN_i(t) - R_i(t) \lambda_{0n}(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} dt \right\} \right) \tag{A.16}$$

$$+ \sqrt{n}\mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right. \tag{A.17}$$

$$\left. \times \hat{\lambda}_0(t, \alpha_n, \beta_n) \left\{ A_i - \bar{A}_n(t, \alpha_n, \beta_n) \right\} dt \right) (\alpha_n - \tilde{\alpha}) \tag{A.18}$$

$$+ \sqrt{n}\mathbb{E}_n \left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{ L_i - \bar{L}_n(t, \alpha_n, \beta_n) \right\} \hat{\lambda}_0(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} \right. \tag{A.19}$$

$$\times \left\{A_i - \bar{A}_n(t,\alpha_n,\beta_n)\right\} dt\Big] (\alpha_n - \tilde{\alpha}) \tag{A.20}$$

$$+ o_{P_n}(1). \tag{A.21}$$

Note that the sequence $\sqrt{n}(\alpha_n - \tilde{\alpha})$ is multiplied by

$$\mathbb{E}_n \Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i}$$

$$\times \hat{\lambda}_0(t,\alpha_n,\beta_n) \left\{A_i - \bar{A}_n(t,\alpha_n,\beta_n)\right\} dt\Bigg) \tag{A.22}$$

$$+ \mathbb{E}_n \Bigg[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\} \hat{\lambda}_0(t,\alpha_n,\beta_n) e^{\alpha_n A_i + \beta_n' L_i}$$

$$\times \left\{A_i - \bar{A}_n(t,\alpha_n,\beta_n)\right\} dt\Bigg]. \tag{A.23}$$

By Hölder's inequality, term (A.23) yields

$$\left|\mathbb{E}_n \left[\int_0^\tau R_i(t)(\gamma_n - \hat{\gamma})' \left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\} \hat{\lambda}_0(t,\alpha_n,\beta_n) e^{\alpha_n A_i + \beta_n' L_i} \left\{A_i - \bar{A}_n(t,\alpha_n,\beta_n)\right\} dt\right]\right|$$

$$\leqslant \left|\left|\mathbb{E}_n \left[\int_0^\tau R_i(t) \left\{L_i - \bar{L}_n(t,\alpha_n,\beta_n)\right\} \hat{\lambda}_0(t,\alpha_n,\beta_n) e^{\alpha_n A_i + \beta_n' L_i} \left\{A_i - \bar{A}_n(t,\alpha_n,\beta_n)\right\} dt\right]\right|\right|_\infty$$

$$\times ||\gamma_n - \hat{\gamma}||_1$$

$$= o_{P_n}(1),$$

where the last equation follows from the ultra-sparsity condition, Assumption 4.3 and the fact that the L_∞-norm is bounded by a positive constant because of Assumption 1, Assumption 2.1 and Assumption 3. For term (A.22), we have

$$\mathbb{E}_n \Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t,\alpha_n,\beta_n)$$

$$\times \left\{A_i - \bar{A}_n(t,\alpha_n,\beta_n)\right\} dt\Bigg)$$

$$= \mathbb{E}_n \Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \lambda_{0n}(t,\alpha_n,\beta_n)$$

$$\times \left\{A_i - \bar{A}^*(t,\alpha_n,\beta_n)\right\} dt\Bigg) \tag{A.24}$$

$$+ \mathbb{E}_n \Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i}$$

$$\times \left\{\hat{\lambda}_0(t,\alpha_n,\beta_n) - \lambda_{0n}(t,\alpha_n,\beta_n)\right\} \left\{A_i - \bar{A}_n(t,\alpha_n,\beta_n)\right\} dt\Bigg) \tag{A.25}$$

$$+\,\mathbb{E}_n\left(\int_0^{\tau}\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\lambda_{0n}(t,\alpha_n,\beta_n)\right.$$
$$\left.\times\left\{\bar{A}^*(t,\alpha_n,\beta_n)-\bar{A}_n(t,\alpha_n,\beta_n)\right\}dt\right). \tag{A.26}$$

For term (A.25), we have

$$\left|\mathbb{E}_n\left(\int_0^{\tau}\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\frac{\mathbb{E}_n\left\{dM(t)\right\}}{\mathbb{E}_n\left\{R(t)e^{\alpha_n A+\beta_n' L}\right\}}\right.\right.$$
$$\left.\left.\times\left\{A_i-\bar{A}_n(t,\alpha_n,\beta_n)\right\}dt\right)\right|$$
$$\leqslant\sup_{t\in[0,\tau]}\left|\mathbb{E}_n\left(\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\right.\right.$$
$$\left.\left.\times\left\{A_i-\bar{A}_n(t,\alpha_n,\beta_n)\right\}\right)\right|\int_0^{\tau}\left|\frac{\mathbb{E}_n\left\{dM(t)\right\}}{\mathbb{E}_n\left\{R(t)e^{\alpha_n A+\beta_n' L}\right\}}\right|$$
$$=o_{P_n}(1),$$

as the supremum is bounded by a positive constant, $\int_0^{\tau}\left|\frac{1}{n}\sum_{j=1}^{n}dM_j(t)\right| = O_{P_n}(n^{-1/2})$ (Fang et al., 2017) and because of Assumption 3.

For term (A.26), we have

$$\left|\mathbb{E}_n\left(\int_0^{\tau}\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\lambda_{0n}(t,\alpha_n,\beta_n)\right.\right.$$
$$\left.\left.\times\left\{\bar{A}^*(t,\alpha_n,\beta_n)-\bar{A}_n(t,\alpha_n,\beta_n)\right\}dt\right)\right|$$
$$\leqslant\tau\sup_{t\in[0,\tau]}\left|\mathbb{E}_n\left(\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\lambda_{0n}(t,\alpha_n,\beta_n)\right)\right|\cdot$$
$$\times\sup_{t\in[0,\tau]}\left|\bar{A}^*(t,\alpha_n,\beta_n)-\bar{A}_n(t,\alpha_n,\beta_n)\right|$$
$$=O_{P_n}\left(n^{-1/2}\right)\sup_{t\in[0,\tau]}\left|\bar{A}^*(t,\alpha_n,\beta_n)-\bar{A}_n(t,\alpha_n,\beta_n)\right|,$$

where the last equality follows from Assumption 2, the uniform weak law of large numbers and definition of a supremum (see term R_1 for a similar reasoning). Under the ultra-sparsity condition and Assumption 7, this term converges in probability to zero.

By the weak law of large numbers, term (A.24) converges in probability to

$$\mathbb{E}_{P_n}\left(\int_0^{\tau}\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\lambda_{0n}(t,\alpha_n,\beta_n)\right.$$

$$\times \left\{A_i - \bar{A}^*(t, \alpha_n, \beta_n)\right\} dt \Big) .$$

Equation (A.15) can then be rewritten as

$$\begin{aligned}
0 = &\sqrt{n}\mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t, \alpha_n, \beta_n)\right\}\right] \right. \\
&\left. \times \left\{ dN_i(t) - R_i(t)\lambda_{0n}(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} dt \right\} \right) \\
&+ \sqrt{n} \left[\mathbb{E}_{P_n} \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t, \alpha_n, \beta_n)\right\}\right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \lambda_{0n}(t, \alpha_n, \beta_n) \right. \right. \\
&\qquad \left. \left. \times \left\{A_i - \bar{A}^*(t, \alpha_n, \beta_n)\right\} dt \right) + o_{P_n}(1) \right] (\alpha_n - \tilde{\alpha}) + o_{P_n}(1).
\end{aligned}$$

Rewriting this gives

$$\begin{aligned}
\sqrt{n}(\tilde{\alpha} - \alpha_n) = &\mathbb{E}_{P_n} \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t, \alpha_n, \beta_n)\right\}\right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \right. \\
&\left. \times \lambda_{0n}(t, \alpha_n, \beta_n) \left\{A_i - \bar{A}^*(t, \alpha_n, \beta_n)\right\} dt \right)^{-1} \\
&\times \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t, \alpha_n, \beta_n)\right\}\right] \\
&\times \left\{ dN_i(t) - R_i(t)\lambda_{0n}(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} dt \right\} \qquad \text{(A.27)} \\
&+ o_{P_n}(1).
\end{aligned}$$

To obtain the asymptotic distribution of $\sqrt{n}(\tilde{\alpha} - \alpha_n)$, we verify the Lyapunov condition. Indeed, as Expression (A.27) has mean zero under a correctly specified Cox model for survival time T, and under the assumptions $V \geqslant c > 0$,

$$\mathbb{E}_n \left[\left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t, \alpha_n, \beta_n)\right\}\right] dM_i(t) \right)^2 \right] = O(1),$$

and

$$\mathbb{E}_{P_n} \left[\left| \int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{L_i - \bar{L}^*(t, \alpha_n, \beta_n)\right\}\right] dM_i(t) \right|^q \right] < C$$

for $\epsilon > 0$ such that $4 + 2\epsilon \leqslant q$ with $4 < q < \infty$ and for $0 < c < C < \infty$, the Lyapunov condition holds. By the Lyapunov central limit theorem, $\sqrt{n}(\tilde{\alpha} - \alpha_n) \xrightarrow{d} N(0, \Sigma_n^2)$ follows.

Part 2: Estimation of Variance

In this part, we prove the consistency of the estimator $\widetilde{\Sigma}_n^2$ of the variance, where

$$\begin{aligned}\widetilde{\Sigma}_n^2 =&\widetilde{V}^{-1}\mathbb{E}_n\left[\left(\int_0^\tau \left[A_i - \bar{A}_n(t,\tilde{\alpha},\hat{\beta}) - \hat{\gamma}'\left\{L_i - \bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}\right]\right.\right.\\ &\left.\left.\times\left\{dN_i(t) - R_i(t)e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})\right\}\right)^2\right]\widetilde{V}^{-1},\end{aligned}$$

with

$$\begin{aligned}\widetilde{V} =&\mathbb{E}_n\left(\int_0^\tau \left[A_i - \bar{A}_n(t,\tilde{\alpha},\hat{\beta}) - \hat{\gamma}'\left\{L_i - \bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}\right] R_i(t)e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})\right.\\ &\left.\times\left\{A_i - \bar{A}_n(t,\tilde{\alpha},\hat{\beta})\right\}dt\right).\end{aligned}$$

First, we prove that

$$\mathbb{E}_n\left(\int_0^\tau \left[A_i - \bar{A}_n(t,\tilde{\alpha},\hat{\beta}) - \hat{\gamma}'\left\{L_i - \bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}\right] R_i(t)e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})\left\{A_i - \bar{A}_n(t,\tilde{\alpha},\hat{\beta})\right\}dt\right)^{-1}$$

converges in probability to

$$\begin{aligned}&\mathbb{E}_{P_n}\left(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\lambda_{0n}(t,\alpha_n,\beta_n)\right.\\ &\left.\times\left\{A_i - \bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\right)^{-1}.\end{aligned}$$

Define

$$\begin{aligned}\hat{W}_i(\alpha,\beta,\gamma) =& \int_0^\tau \left[A_i - \bar{A}_n(t,\alpha,\beta) - \gamma'\left\{L_i - \bar{L}_n(t,\alpha,\beta)\right\}\right] R_i(t)e^{\alpha A_i+\beta' L_i}\hat{\lambda}_0(t,\alpha,\beta)\\ &\times\left\{A_i - \bar{A}_n(t,\alpha,\beta)\right\}dt\end{aligned}$$

and its corresponding population value

$$\begin{aligned}W_i^*(\alpha,\beta,\gamma) =& \int_0^\tau \left[A_i - \bar{A}^*(t,\alpha,\beta) - \gamma'\left\{L_i - \bar{L}^*(t,\alpha,\beta)\right\}\right] R_i(t)e^{\alpha A_i+\beta' L_i}\lambda_{0n}(t,\alpha,\beta)\\ &\times\left\{A_i - \bar{A}^*(t,\alpha,\beta)\right\}dt.\end{aligned}$$

Step 1

We first prove that

$$\mathbb{E}_n\{W_i^*(\alpha_n,\beta_n,\gamma_n)\} = \mathbb{E}_{P_n}\{W_i^*(\alpha_n,\beta_n,\gamma_n)\} + o_{P_n}(1). \tag{A.28}$$

Consider,

$$P_{P_n}\left[|\mathbb{E}_n\{W_i^*(\alpha_n,\beta_n,\gamma_n)\} - \mathbb{E}_{P_n}\{W_i^*(\alpha_n,\beta_n,\gamma_n)\}| > \epsilon\right]$$

$$
\begin{aligned}
&\leqslant \frac{1}{\epsilon^2}\mathbb{E}_{P_n}\left[\left|\frac{1}{n}\sum_{i=1}^n W_i^*(\alpha_n,\beta_n,\gamma_n)-\mathbb{E}_{P_n}\{W_i^*(\alpha_n,\beta_n,\gamma_n)\}\right|^2\right]\\
&\leqslant \frac{1}{\epsilon^2 n^2}\left(2-\frac{1}{n}\right)\sum_{i=1}^n \mathbb{E}_{P_n}\left[\left|W_i^*(\alpha_n,\beta_n,\gamma_n)-\mathbb{E}_{P_n}\{W_i^*(\alpha_n,\beta_n,\gamma_n)\}\right|^2\right],
\end{aligned}
$$

where we consecutively apply Chebyshev's inequaltiy and the Von Bahr-Esseen inequality (von Bahr and Esseen, 1965). The last one states that, for independent mean zero variables $X_1,\dots,X_n$ and $q\in[1,2]$, we have

$$
E\left(\left|\sum_{i=1}^n X_i\right|^q\right)\leqslant\left(2-\frac{1}{n}\right)\sum_{i=1}^n E\left(|X_i|^q\right).
$$

Under the assumption that

$$
\begin{aligned}
&Var_{P_n}\left(\int_0^\tau \left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\lambda_{0n}(t,\alpha_n,\beta_n)\right.\\
&\qquad\left.\times\left\{A_i-\bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\right)=O(1),
\end{aligned}
$$

result (A.28) follows from

$$
\mathbb{E}_{P_n}\left[\left|W_i^*(\alpha_n,\beta_n,\gamma_n)-\mathbb{E}_{P_n}\{W_i^*(\alpha_n,\beta_n,\gamma_n)\}\right|^2\right]=Var_{P_n}\{W_i^*(\alpha_n,\beta_n,\gamma_n)\}=O(1).
$$

Step 2

In this step we show that

$$
\mathbb{E}_n\{\hat{W}_i(\tilde{\alpha},\hat{\beta},\hat{\gamma})\}=\mathbb{E}_n\{W_i^*(\alpha_n,\beta_n,\gamma_n)\}+o_{P_n}(1). \tag{A.29}
$$

$$
\begin{aligned}
&\mathbb{E}_n\left(\int_0^\tau \left[A_i-\bar{A}_n(t,\tilde{\alpha},\hat{\beta})-\hat{\gamma}'\left\{L_i-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}\right]R_i(t)e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})\right. && \text{(A.30)}\\
&\left.\times\left\{A_i-\bar{A}_n(t,\tilde{\alpha},\hat{\beta})\right\}dt\right)\\
&=\mathbb{E}_n\left(\int_0^\tau \left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\lambda_{0n}(t,\alpha_n,\beta_n)\right. && \text{(A.31)}\\
&\left.\times\left\{A_i-\bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\right)\\
&+\mathbb{E}_n\left(\int_0^\tau \left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\right.\\
&\left.\times\{\hat{\lambda}_0(t,\alpha_n,\beta_n)-\lambda_{0n}(t,\alpha_n,\beta_n)\}\left\{A_i-\bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\right) && \text{(A.32)}
\end{aligned}
$$

$$+\mathbb{E}_n\left(\int_0^\tau \left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)A_i\right.$$
$$\left.\times\left\{\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}-\hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_nA_i+\beta_n'L_i}\right\}dt\right) \tag{A.33}$$

$$-\mathbb{E}_n\left(\int_0^\tau \left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)\right.$$
$$\left.\times\left\{\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\bar{A}_n(t,\tilde{\alpha},\hat{\beta})-\hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_nA_i+\beta_n'L_i}\bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\right) \tag{A.34}$$
$$+\mathbb{E}_n\left[\int_0^\tau\{\bar{A}^*(t,\alpha_n,\beta_n)-\bar{A}_n(t,\tilde{\alpha},\hat{\beta})\}R_i(t)e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})\left\{A_i-\bar{A}_n(t,\tilde{\alpha},\hat{\beta})\right\}dt\right] \tag{A.35}$$
$$-\mathbb{E}_n\left[\int_0^\tau\gamma_n'\{\bar{L}^*(t,\alpha_n,\beta_n)-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\}R_i(t)e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})\right.$$
$$\left.\times\left\{A_i-\bar{A}_n(t,\tilde{\alpha},\hat{\beta})\right\}dt\right] \tag{A.36}$$
$$+\mathbb{E}_n\left[\int_0^\tau(\gamma_n-\hat{\gamma})'\left\{L_i-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right.$$
$$\left.\times\left\{A_i-\bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\right] \tag{A.37}$$
$$+\mathbb{E}_n\left[\int_0^\tau(\gamma_n-\hat{\gamma})'\left\{L_i-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}R_i(t)A_i\right.$$
$$\left.\times\left\{e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})-e^{\alpha_nA_i+\beta_n'L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right\}dt\right] \tag{A.38}$$
$$-\mathbb{E}_n\left[\int_0^\tau(\gamma_n-\hat{\gamma})'R_i(t)\left\{L_i-\bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}\right.$$
$$\left.\times\left\{\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\bar{A}_n(t,\tilde{\alpha},\hat{\beta})-\hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_nA_i+\beta_n'L_i}\bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\right]. \tag{A.39}$$

First, note that following a similar reasoning as for term R_1, it holds for term (A.32) that

$$\left|\mathbb{E}_n\left(\int_0^\tau \left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\right.\right.$$
$$\left.\left.\times\{\hat{\lambda}_0(t,\alpha_n,\beta_n)-\lambda_{0n}(t,\alpha_n,\beta_n)\}\left\{A_i-\bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\right)\right|$$
$$\leqslant\delta^{-1}\sup_{t\in[0,\tau]}\left|\mathbb{E}_n\left(\left[A_i-\bar{A}^*(t,\alpha_n,\beta_n)-\gamma_n'\left\{L_i-\bar{L}^*(t,\alpha_n,\beta_n)\right\}\right]R_i(t)e^{\alpha_nA_i+\beta_n'L_i}\right.\right.$$
$$\left.\left.\times\left\{A_i-\bar{A}^*(t,\alpha_n,\beta_n)\right\}\right)\right|\int_0^\tau\left|\mathbb{E}_n\left\{dM_j(t)\right\}\right|.$$

By the fact that the supremum can be bounded by a constant (see Assumption 1 and Assumption

2) and $\int_0^\tau \left|\frac{1}{n}\sum_{j=1}^n dM_j(t)\right| = O_{P_n}(n^{-1/2})$, it follows that term (A.32) converges in probability to zero.

Next, consider term (A.33),

$$\begin{aligned}
&\Bigg|\mathbb{E}_n\Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t)A_i \\
&\times\left\{\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i} - \hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i}\right\}dt\Bigg)\Bigg| \\
&\leqslant \Bigg|\Bigg|\mathbb{E}_n\Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t)A_i\hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i} \\
&\times\left\{X_i - \bar{X}_n(t,\alpha_n,\beta_n)\right\}\Bigg)\Bigg|\Bigg|_\infty \cdot \left|\left|\theta_n - \tilde{\theta}\right|\right|_1 + O_{P_n}(||\theta_n - \tilde{\theta}||_2^2),
\end{aligned}$$

by following a Taylor expansion and using Hölder's inequality. From Assumption 4.1, Assumption 4.2, the consistency of $\tilde{\alpha}$ and the fact that the L_∞-norm of the population mean,

$$\begin{aligned}
&\mathbb{E}_n\Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t)A_i\hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i} \\
&\quad\times\left\{X_i - \bar{X}_n(t,\alpha_n,\beta_n)\right\}\Bigg),
\end{aligned}$$

can be bounded by a constant (see Assumptions 1-3), it follows that term (A.33) converges in probability to zero. Similarly, for term (A.34),

$$\begin{aligned}
&\Bigg|\mathbb{E}_n\Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t) \\
&\times\ \left\{\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\bar{A}_n(t,\hat{\alpha},\hat{\beta}) - \hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i}\bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\Bigg)\Bigg| \\
&\leqslant \Bigg|\mathbb{E}_n\Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t) \\
&\times\ \left\{\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta})e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\bar{A}_n(t,\hat{\alpha},\hat{\beta}) - \hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i}\bar{A}_n(t,\alpha_n,\beta_n)\right\}dt\Bigg)\Bigg| \\
&+\Bigg|\mathbb{E}_n\Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t)\hat{\lambda}_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i} \\
&\times\ \left\{\bar{A}_n(t,\hat{\alpha},\hat{\beta}) - \bar{A}^*(t,\alpha_n,\beta_n)\right\}dt\Bigg)\Bigg|.
\end{aligned}$$

Here, by following a Taylor expansion and using Hölder's inequality the first term can be bounded by

$$\Bigg|\Bigg|\mathbb{E}_n\Bigg(\int_0^\tau \left[A_i - \bar{A}^*(t,\alpha_n,\beta_n) - \gamma_n'\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\}\right] R_i(t)$$

$$\times \frac{\partial}{\partial \theta} \left\{ \hat{\lambda}_0(t, \alpha, \beta) e^{\alpha A_i + \beta' L_i} \bar{A}_n(t, \alpha, \beta) \right\} \Big|_{\theta = \theta_n} \Big) \Big|\Big|_\infty \Big|\Big| \theta_n - \tilde{\theta} \Big|\Big|_1 + O_{P_n}(||\theta_n - \tilde{\theta}||_2^2).$$

From Assumption 4.1, Assumption 4.2, the consistency of $\tilde{\alpha}$ and the fact that the L_∞-norm can be bounded by a constant (see Assumptions 1-3), it follows that term (A.34) converges in probability to zero. The other term can be bounded by

$$\tau \sup_{t \in [0,\tau]} \Big| \mathbb{E}_n \Big(\left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \Big) \Big| \sup_{t \in [0,\tau]} \hat{\lambda}_0(t, \alpha_n, \beta_n)$$
$$\times \sup_{t \in [0,\tau]} \Big| \Big\{ \bar{A}_n(t, \hat{\alpha}, \hat{\beta}) - \bar{A}^*(t, \alpha_n, \beta_n) \Big\} \Big|,$$

which converges, under the ultra-sparsity condition, in probability to zero as the first supremum is bounded by $O_{P_n}(n^{-1/2})$ (following a similar reasoning as for R_1), $\sup_{t\in[0,\tau]} \hat{\lambda}_0(t, \alpha_n, \beta_n)$ is bounded by a constant (see Assumption 3) and the last supremum is bounded by $O_{P_n}\left(\sqrt{\frac{\log(p \vee n)}{n}}\right)$ (see Assumption 7).

Note that term (A.35) and term (A.36) equal 0. Next, by Hölder's inequality for term (A.37) it holds that

$$\Big| \mathbb{E}_n \Big[\int_0^\tau (\gamma_n - \hat{\gamma})' \Big\{ L_i - \bar{L}_n(t, \tilde{\alpha}, \hat{\beta}) \Big\} R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t, \alpha_n, \beta_n) \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} dt \Big] \Big|$$
$$\leqslant \Big| \mathbb{E}_n \Big[\int_0^\tau (\gamma_n - \hat{\gamma})' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t, \alpha_n, \beta_n) \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} dt \Big] \Big|$$
$$+ \Big| \mathbb{E}_n \Big[\int_0^\tau (\gamma_n - \hat{\gamma})' \Big\{ \bar{L}^*(t, \alpha_n, \beta_n) - \bar{L}_n(t, \tilde{\alpha}, \hat{\beta}) \Big\} R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t, \alpha_n, \beta_n) \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} dt \Big] \Big|$$
$$\leqslant \Big|\Big| \mathbb{E}_n \Big[\int_0^\tau \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t, \alpha_n, \beta_n) \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} dt \Big] \Big|\Big|_\infty$$
$$\times ||\gamma_n - \hat{\gamma}||_1$$
$$+ \Big|\Big| \mathbb{E}_n \Big[\int_0^\tau \Big\{ \bar{L}^*(t, \alpha_n, \beta_n) - \bar{L}_n(t, \tilde{\alpha}, \hat{\beta}) \Big\} R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t, \alpha_n, \beta_n) \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} dt \Big] \Big|\Big|_\infty$$
$$\times ||\gamma_n - \hat{\gamma}||_1$$
$$\leqslant \tau \sup_{t \in [0,\tau]} \Big|\Big| \mathbb{E}_n \Big[\left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t, \alpha_n, \beta_n) \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} \Big] \Big|\Big|_\infty$$
$$\times ||\gamma_n - \hat{\gamma}||_1$$
$$+ \tau \sup_{t \in [0,\tau]} \Big| \mathbb{E}_n \Big[R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \hat{\lambda}_0(t, \alpha_n, \beta_n) \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} \Big] \Big|$$

$$\times \sup_{t\in[0,\tau]} \left|\left|\bar{L}^*(t,\alpha_n,\beta_n) - \bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right|\right|_\infty ||\gamma_n - \hat{\gamma}||_1$$

$$= o_{P_n}(1),$$

where the last equality follows from Assumption 4.3, the fact that

$$\left|\left|\mathbb{E}_n\left[\left\{L_i - \bar{L}^*(t,\alpha_n,\beta_n)\right\} R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\left\{A_i - \bar{A}^*(t,\alpha_n,\beta_n)\right\}\right]\right|\right|_\infty$$

and

$$\left|\mathbb{E}_n\left[R_i(t)e^{\alpha_n A_i+\beta_n' L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\left\{A_i - \bar{A}^*(t,\alpha_n,\beta_n)\right\}\right]\right|$$

are bounded for all $t \in [0,\tau]$ (see Assumption 1-3) and the fact that for all $t \in [0,\tau]$

$$\begin{aligned}\left|\left|\bar{L}^*(t,\alpha_n,\beta_n) - \bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right|\right|_\infty &\leqslant \left|\left|\bar{L}^*(t,\alpha_n,\beta_n) - \bar{L}_n(t,\alpha_n,\beta_n)\right|\right|_\infty \\ &+ \left|\left|\bar{L}_n(t,\alpha_n,\beta_n) - \bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right|\right|_\infty \\ &\leqslant O_{P_n}\left(s_\theta\sqrt{\frac{\log(p\vee n)}{n}}\right),\end{aligned}$$

by Assumption 7 and Lemma E.2 in Appendix of Fang et al. (2017), which relies on primitive conditions required for uniform consistency of Lasso estimators and Assumption 1 and Assumption 2.

Next, by Hölder's inequality, term (A.38) yields

$$\begin{aligned}&\left|\mathbb{E}_n\left[\int_0^\tau (\gamma_n-\hat{\gamma})'\left\{L_i - \bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}R_i(t)A_i\left\{e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta}) - e^{\alpha_n A_i+\beta_n' L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right\}dt\right]\right| \\ &\leqslant \left|\left|\mathbb{E}_n\left[\int_0^\tau L_i R_i(t)A_i\left\{e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta}) - e^{\alpha_n A_i+\beta_n' L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right\}dt\right]\right|\right|_\infty ||\gamma_n-\hat{\gamma}||_1 \\ &+ \left|\left|\mathbb{E}_n\left[\int_0^\tau \left\{\bar{L}^*(t,\alpha_n,\beta_n) - \bar{L}_n(t,\tilde{\alpha},\hat{\beta})\right\}R_i(t)A_i\left\{e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta}) - e^{\alpha_n A_i+\beta_n' L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right\}dt\right]\right|\right|_\infty \\ &\times ||\gamma_n-\hat{\gamma}||_1 \\ &+ \left|\left|\mathbb{E}_n\left[\int_0^\tau \bar{L}^*(t,\alpha_n,\beta_n)R_i(t)A_i\left\{e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta}) - e^{\alpha_n A_i+\beta_n' L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right\}dt\right]\right|\right|_\infty ||\gamma_n-\hat{\gamma}||_1.\end{aligned}$$

For the first term, we have

$$\left|\left|\mathbb{E}_n\left[\int_0^\tau L_i R_i(t)A_i\left\{e^{\tilde{\alpha}A_i+\hat{\beta}'L_i}\hat{\lambda}_0(t,\tilde{\alpha},\hat{\beta}) - e^{\alpha_n A_i+\beta_n' L_i}\hat{\lambda}_0(t,\alpha_n,\beta_n)\right\}dt\right]\right|\right|_\infty ||\gamma_n-\hat{\gamma}||_1$$

$$= \left|\left|\mathbb{E}_n\left(\int_0^\tau \left[\frac{\mathbb{E}_n\left\{R_j(t)A_jL_je^{\tilde\alpha A_j+\hat\beta' L_j}\right\}}{\mathbb{E}_n\left\{R_j(t)e^{\tilde\alpha A_j+\hat\beta' L_j}\right\}} - \frac{\mathbb{E}_n\left\{R_j(t)A_jL_je^{\alpha_n A_j+\beta_n' L_j}\right\}}{\mathbb{E}_n\left\{R_j(t)e^{\alpha_n A_j+\beta_n' L_j}\right\}}\right] dN_i(t)\right)\right|\right|_\infty$$
$$\times ||\gamma_n - \hat\gamma||_1$$
$$\leqslant \sup_{t\in[0,\tau]}\left|\left|\frac{\mathbb{E}_n\left\{R_j(t)A_jL_je^{\tilde\alpha A_j+\hat\beta' L_j}\right\}}{\mathbb{E}_n\left\{R_j(t)e^{\tilde\alpha A_j+\hat\beta' L_j}\right\}} - \frac{\mathbb{E}_n\left\{R_j(t)A_jL_je^{\alpha_n A_j+\beta_n' L_j}\right\}}{\mathbb{E}_n\left\{R_j(t)e^{\alpha_n A_j+\beta_n' L_j}\right\}}\right|\right|_\infty$$
$$\times \int_0^\tau |\mathbb{E}_n\left\{dN_i(t)\right\}|\, ||\gamma_n - \hat\gamma||_1$$
$$\leqslant \tau \sup_{t\in[0,\tau]}\left|\left|\frac{\mathbb{E}_n\left\{R_j(t)A_jL_je^{\tilde\alpha A_j+\hat\beta' L_j}\right\}}{\mathbb{E}_n\left\{R_j(t)e^{\tilde\alpha A_j+\hat\beta' L_j}\right\}} - \frac{\mathbb{E}_n\left\{R_j(t)A_jL_je^{\alpha_n A_j+\beta_n' L_j}\right\}}{\mathbb{E}_n\left\{R_j(t)e^{\alpha_n A_j+\beta_n' L_j}\right\}}\right|\right|_\infty ||\gamma_n - \hat\gamma||_1\,,$$

which converges in probability to zero by Assumption 4.3 and the fact that the L_∞-norm is bounded by $Cs_\beta\sqrt{\frac{\log(p\vee n)}{n}}$ for a certain constant C, which can be shown by following a similar reasoning as in Lemma E.2 of Fang et al. (2017). The other two terms left for expression (A.38) can be shown to converge in probability to zero in a similar way.

Finally, consider the sample mean in term (A.39)

$$\mathbb{E}_n\left[\int_0^\tau R_i(t)\left\{L_i - \bar L_n(t,\tilde\alpha,\hat\beta)\right\}\left\{\hat\lambda_0(t,\tilde\alpha,\hat\beta)e^{\tilde\alpha A_i+\hat\beta' L_i}\bar A_n(t,\tilde\alpha,\hat\beta)\right.\right.$$
$$\left.\left.-\hat\lambda_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i}\bar A^*(t,\alpha_n,\beta_n)\right\} dt\right]$$
$$= \mathbb{E}_n\left[\int_0^\tau R_i(t)\left\{L_i - \bar L_n(t,\tilde\alpha,\hat\beta)\right\}\hat\lambda_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i}\left\{\bar A_n(t,\tilde\alpha,\hat\beta) - \bar A^*(t,\alpha_n,\beta_n)\right\} dt\right]$$
$$+ \mathbb{E}_n\left[\int_0^\tau R_i(t)\left\{L_i - \bar L_n(t,\tilde\alpha,\hat\beta)\right\}\left\{\hat\lambda_0(t,\tilde\alpha,\hat\beta)e^{\tilde\alpha A_i+\hat\beta' L_i} - \hat\lambda_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i}\right\}\bar A^*(t,\alpha_n,\beta_n)dt\right]$$
$$+ \mathbb{E}_n\left[\int_0^\tau R_i(t)\left\{L_i - \bar L_n(t,\tilde\alpha,\hat\beta)\right\}\left\{\hat\lambda_0(t,\tilde\alpha,\hat\beta)e^{\tilde\alpha A_i+\hat\beta' L_i} - \hat\lambda_0(t,\alpha_n,\beta_n)e^{\alpha_n A_i+\beta_n' L_i}\right\}\right.$$
$$\left.\times\left\{\bar A_n(t,\tilde\alpha,\hat\beta) - \bar A^*(t,\alpha_n,\beta_n)\right\} dt\right].$$

These terms, multiplied by $(\gamma_n - \hat\gamma)$, can be bounded in a similar way as term (A.38). We conclude that,

$$\mathbb{E}_n\left(\int_0^\tau \left[A_i - \bar A_n(t,\tilde\alpha,\hat\beta) - \hat\gamma'\left\{L_i - \bar L_n(t,\tilde\alpha,\hat\beta)\right\}\right]R_i(t)e^{\tilde\alpha A_i+\hat\beta' L_i}\hat\lambda_0(t,\tilde\alpha,\hat\beta)\right.$$
$$\left.\times\left\{A_i - \bar A_n(t,\tilde\alpha,\hat\beta)\right\} dt\right)$$

$$= \mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \lambda_0(t, \alpha_n, \beta_n) \right.$$

$$\left. \times \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} dt \right)$$

$$+ o_{P_n}(1),$$

which proves Equation (A.29). It follows from Step 1 and Step 2 that

$$\mathbb{E}_n\{\hat{W}_i(\tilde{\alpha}, \hat{\beta}, \hat{\gamma})\} = \mathbb{E}_{P_n}\{W_i^*(\alpha_n, \beta_n, \gamma_n)\} + o_{P_n}(1).$$

Step 3

In this step we show that

$$\mathbb{E}_n\{\hat{W}_i(\tilde{\alpha}, \hat{\beta}, \hat{\gamma})\}^{-1} = \mathbb{E}_{P_n}\{W_i^*(\alpha_n, \beta_n, \gamma_n)\}^{-1} + o_{P_n}(1). \tag{A.40}$$

First note that $\mathbb{E}_{P_n}\{W_i^*(\alpha_n, \beta_n, \gamma_n)\}^{-1}$ exists under the assumption that $\forall n : |\mathbb{E}_{P_n}\{W_i^*(\alpha_n, \beta_n, \gamma_n)\}| \geqslant \delta > 0$. This is a weak condition, which is fulfilled if there is some variation in the exposure at some times t and some levels of L. In particular, this translates to an assumption of the invertibility of the relevant components (i.e., corresponding with parameter of interest α) of the information matrix corresponding with the partial likelihood.

Given that $\forall n : |\mathbb{E}_{P_n}\{W_i^*(\alpha_n, \beta_n, \gamma_n)\}| \geqslant \delta > 0$, the inverse transformation is Lipschitz continuous (Van der Vaart, 1998). Then by the Uniform Continuous Mapping Theorem (Kasy, 2019),

$$\left| \mathbb{E}_n\{\hat{W}_i(\tilde{\alpha}, \hat{\beta}, \hat{\gamma})\}^{-1} - \mathbb{E}_{P_n}\{W_i^*(\alpha_n, \beta_n, \gamma_n)\}^{-1} \right| \xrightarrow{P} 0.$$

To complete this proof for the consistency of $\widetilde{\Sigma}_n^2$, we still need to prove that

$$\mathbb{E}_n \left\{ \left(\int_0^\tau \left[A_i - \bar{A}_n(t, \tilde{\alpha}, \hat{\beta}) - \hat{\gamma}' \left\{ L_i - \bar{L}_n(t, \tilde{\alpha}, \hat{\beta}) \right\} \right] \left\{ dN_i(t) - R_i(t) e^{\tilde{\alpha} A_i + \hat{\beta}' L_i} \hat{\lambda}_0(t, \tilde{\alpha}, \hat{\beta}) \right\} \right)^2 \right\}$$

converges in probability to

$$\mathbb{E}_{P_n} \left\{ \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] dM_i(t) \right)^2 \right\}.$$

This can be shown by essentially repeating Steps 2 and 3 from the proof of Theorem 3.3 in Dukes et al. (2020), which are similar to Steps 1 and 2 of the previous part of our proof.

Part 3: Link to Triple Selection

So far, we have focused on the one-step estimator $\tilde{\alpha}$ based on solving the decorrelated score function. Now, we will link the triple selection procedure with the orthogonal (i.e., debiased) score for which we obtained the former results. Recall that B is the union of the variables estimated to have non-zero coefficients in the first three (selection) steps. Now, define $X_B = (A, L_B')'$, where L_B includes all variables L_k for which $k \in B$. By the last step of the triple selection strategy, we have

$$\mathbb{E}_n \left(\int_0^\tau \left[X_{Bi} - \frac{\mathbb{E}_n \left\{ R_j(t) X_{Bj} e^{\check{\alpha} A_j + \sum_{k\in B} \check{\beta}_k L_{kj}} \right\}}{\mathbb{E}_n \left\{ R_j(t) e^{\check{\alpha} A_j + \sum_{k\in B} \check{\beta}_k L_{kj}} \right\}} \right] dN_i(t) \right) = 0. \tag{A.41}$$

Next, define $\check{\gamma}$ as the solution to

$$\mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}_n(t, \check{\alpha}, \check{\beta}) - \gamma' \left\{ L_i - \bar{L}_n(t, \check{\alpha}, \check{\beta}) \right\} \right] \left\{ L_i - \bar{L}_n(t, \check{\alpha}, \check{\beta}) \right\} dN_i(t) \right) = 0,$$

subject to support$(\gamma) \subseteq B$. Note that by (A.41), taking the linear combination $(1, -\check{\gamma})$, we have

$$\begin{aligned}
0 &= \mathbb{E}_n \left[\int_0^\tau \left\{ A_i - \bar{A}_n(t, \check{\alpha}, \check{\beta}) \right\} dN_i(t) \right] - \check{\gamma}' \mathbb{E}_n \left[\int_0^\tau \left\{ L_i - \bar{L}_n(t, \check{\alpha}, \check{\beta}) \right\} dN_i(t) \right] \\
&= \mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}_n(t, \check{\alpha}, \check{\beta}) - \check{\gamma}' \left\{ L_i - \bar{L}_n(t, \check{\alpha}, \check{\beta}) \right\} \right] dN_i(t) \right) \\
&= \mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}_n(t, \check{\alpha}, \check{\beta}) - \check{\gamma}' \left\{ L_i - \bar{L}_n(t, \check{\alpha}, \check{\beta}) \right\} \right] \left\{ dN_i(t) - R_i(t) e^{\check{\alpha} A_i + \check{\beta}' L_i} \hat{\lambda}_0(t, \check{\alpha}, \check{\beta}) \right\} \right),
\end{aligned}$$

where the last equation follows from

$$\mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}_n(t, \check{\alpha}, \check{\beta}) - \check{\gamma}' \left\{ L_i - \bar{L}_n(t, \check{\alpha}, \check{\beta}) \right\} \right] R_i(t) e^{\check{\alpha} A_i + \check{\beta}' L_i} \hat{\lambda}_0(t, \check{\alpha}, \check{\beta}) \right) = 0.$$

Moreover, since support$(\hat{\gamma}) \subseteq B$ and support$(\hat{\beta}) \subseteq B$, it follows from Part 1 that,

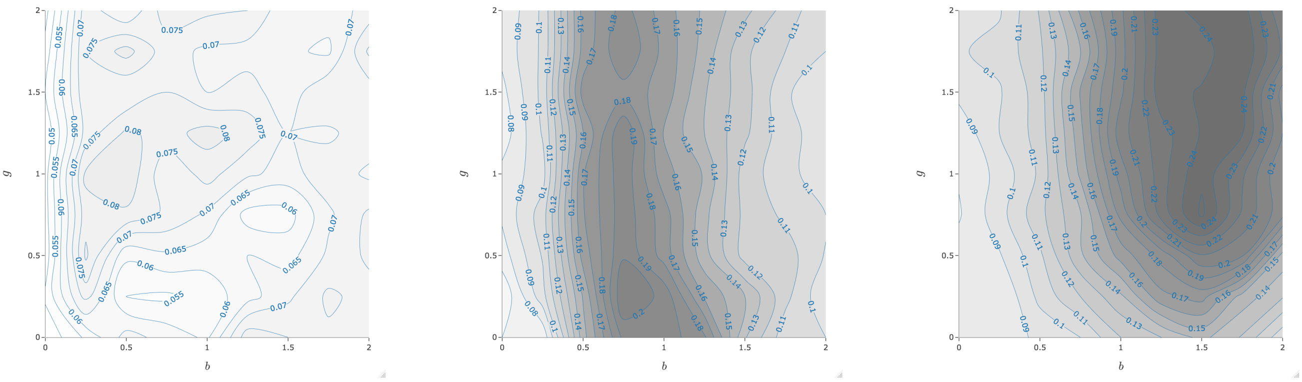
$$\begin{aligned}
&\sqrt{n} \mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}_n(t, \check{\alpha}, \check{\beta}) - \check{\gamma}' \left\{ L_i - \bar{L}_n(t, \check{\alpha}, \check{\beta}) \right\} \right] R_i(t) \left\{ dN_i(t) - e^{\check{\alpha} A_i + \check{\beta}' L_i} \hat{\lambda}_0(t, \check{\alpha}, \check{\beta}) \right\} \right) \\
&= \sqrt{n} \mathbb{E}_n \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] \right. \\
&\left. \times \left\{ dN_i(t) - R_i(t) \lambda_{0n}(t, \alpha_n, \beta_n) e^{\alpha_n A_i + \beta_n' L_i} dt \right\} \right) \\
&+ \sqrt{n} \left\{ \mathbb{E}_{P_n} \left(\int_0^\tau \left[A_i - \bar{A}^*(t, \alpha_n, \beta_n) - \gamma_n' \left\{ L_i - \bar{L}^*(t, \alpha_n, \beta_n) \right\} \right] R_i(t) e^{\alpha_n A_i + \beta_n' L_i} \lambda_{0n}(t, \alpha_n, \beta_n) \right. \right. \\
&\left. \left. \times \left\{ A_i - \bar{A}^*(t, \alpha_n, \beta_n) \right\} dt \right) + o_{P_n}(1) \right\} (\alpha_n - \check{\alpha}) + o_{P_n}(1),
\end{aligned}$$

and consequently that

$$\sqrt{n}\,(\check{\alpha} - \alpha_n) = Z_n + o_{P_n}(1), \quad Z_n \xrightarrow{d} N(0, \Sigma_n^2), \tag{A.42}$$

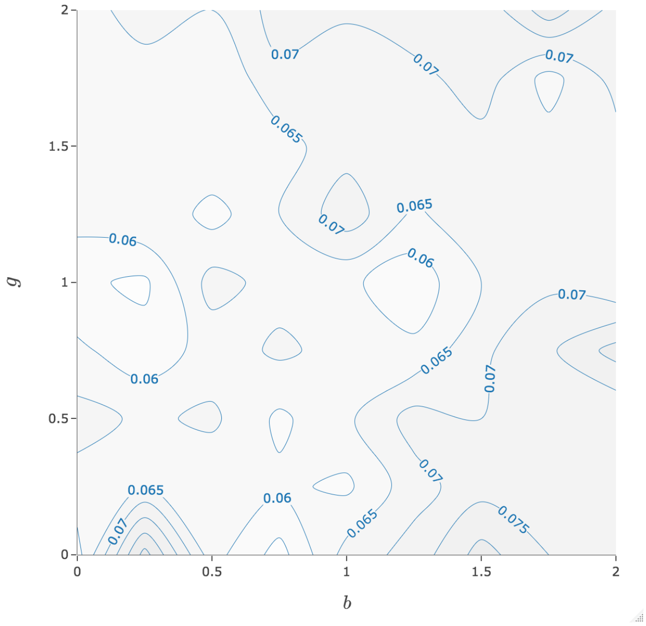
as $n \to \infty$. Finally, since support$(\hat{\gamma}) \subseteq B$ and support$(\hat{\beta}) \subseteq B$, $\widehat{\Sigma}^2$ is a consistent estimator of Σ_n^2 by Part 2 of this proof.

Appendix B: Figures

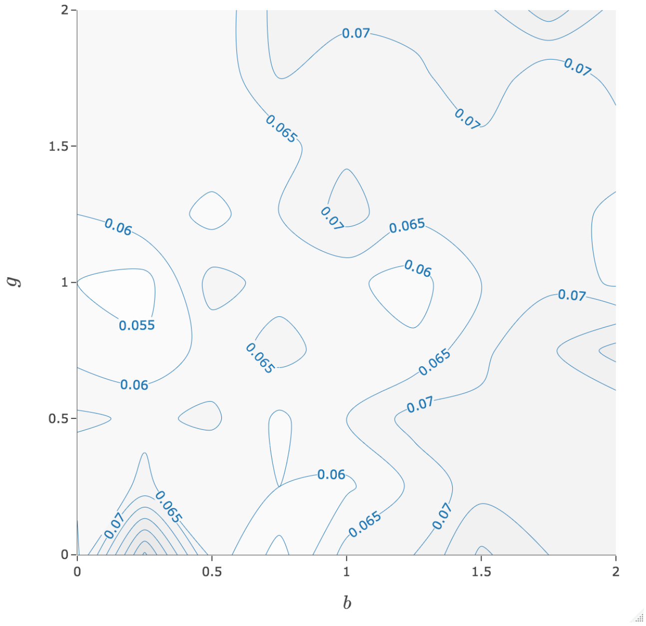


(a) Setting 1(a) with $\rho = 0.50$ (b) Setting 1(b) with $c_A = 1$ (c) Setting 1(b) with $c_A = 2$

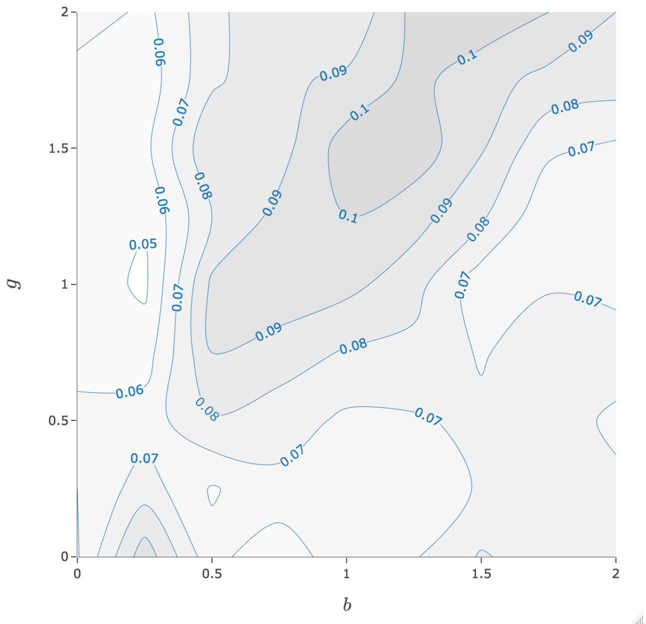
Figure 1: Empirical Type I error rate at the 5% significance level of the double selection approach (i.e., as described in Section 3.2 of the main paper without the censoring model) under Setting 1 with $n = 400$, $p = 30$, $\eta_1 = 1$, $\beta_0 = 0$ and $\gamma_0 = 0$ for different distributions of X.



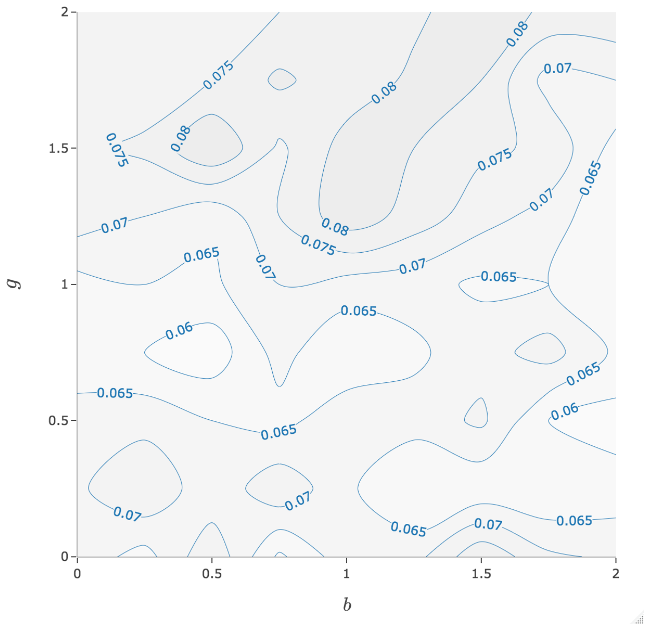
(a) Triple selection



(b) Poor man's approach

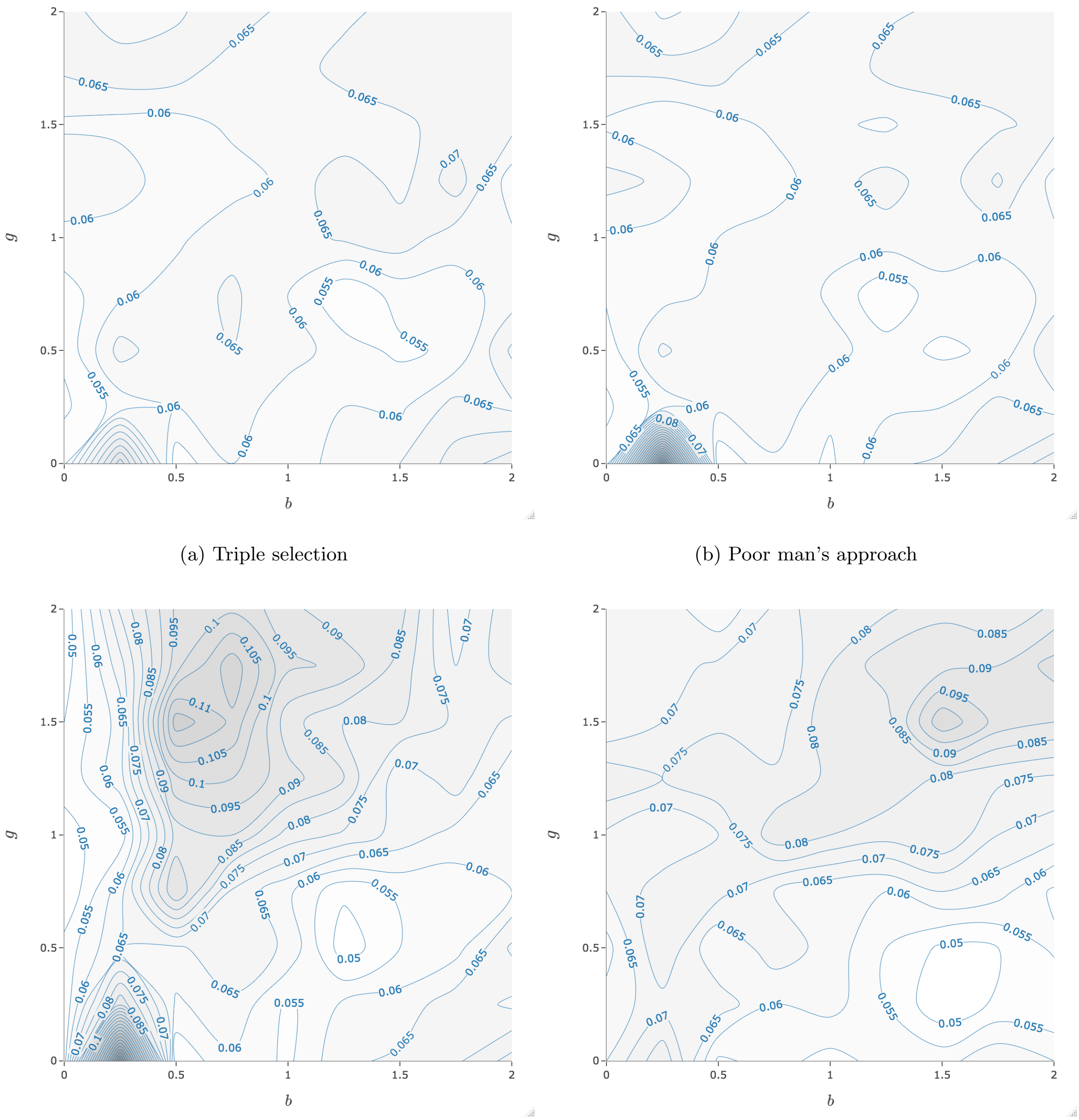


(c) Post-Lasso



(d) Decorrelated test in Fang et al. (2017)

Figure 2: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(a) with $n = 400$, $p = 30$, $\rho = 0.25$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



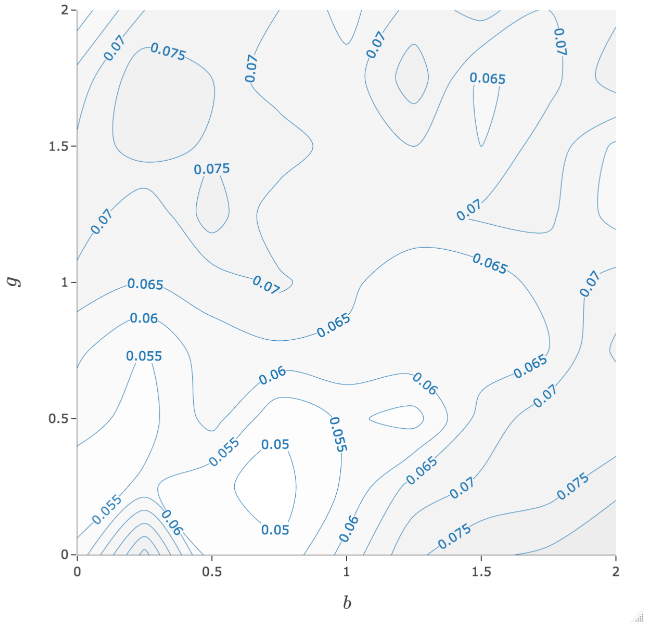
(a) Triple selection

(b) Poor man's approach

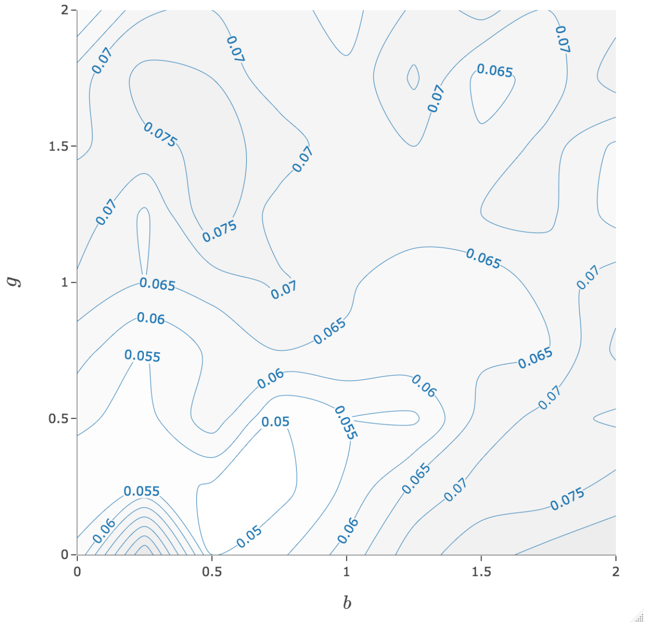
(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

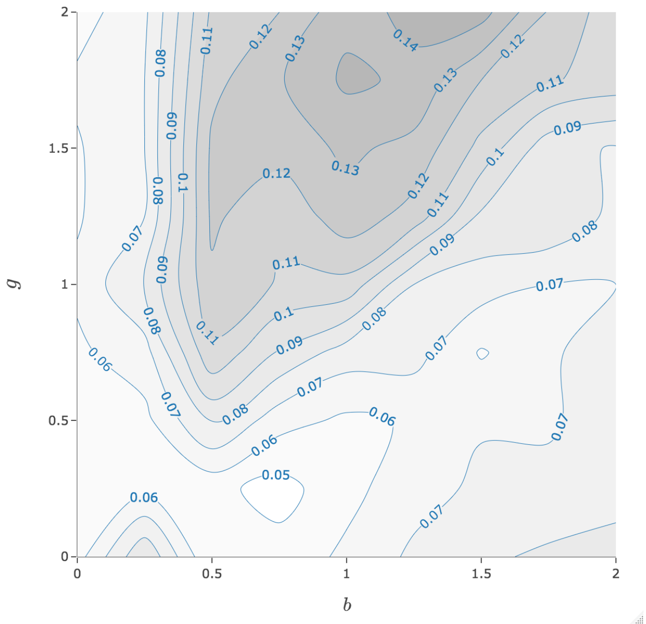
Figure 3: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(a) with $n = 400$, $p = 30$, $\rho = 0.50$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



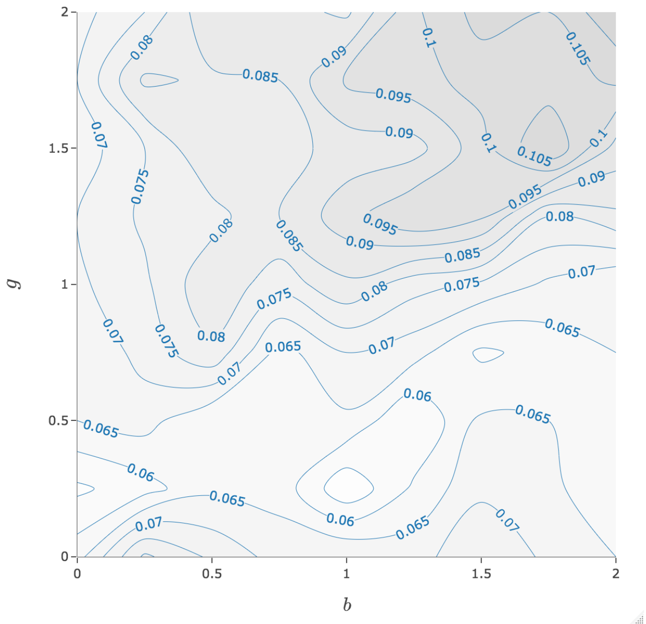
(a) Triple selection



(b) Poor man's approach

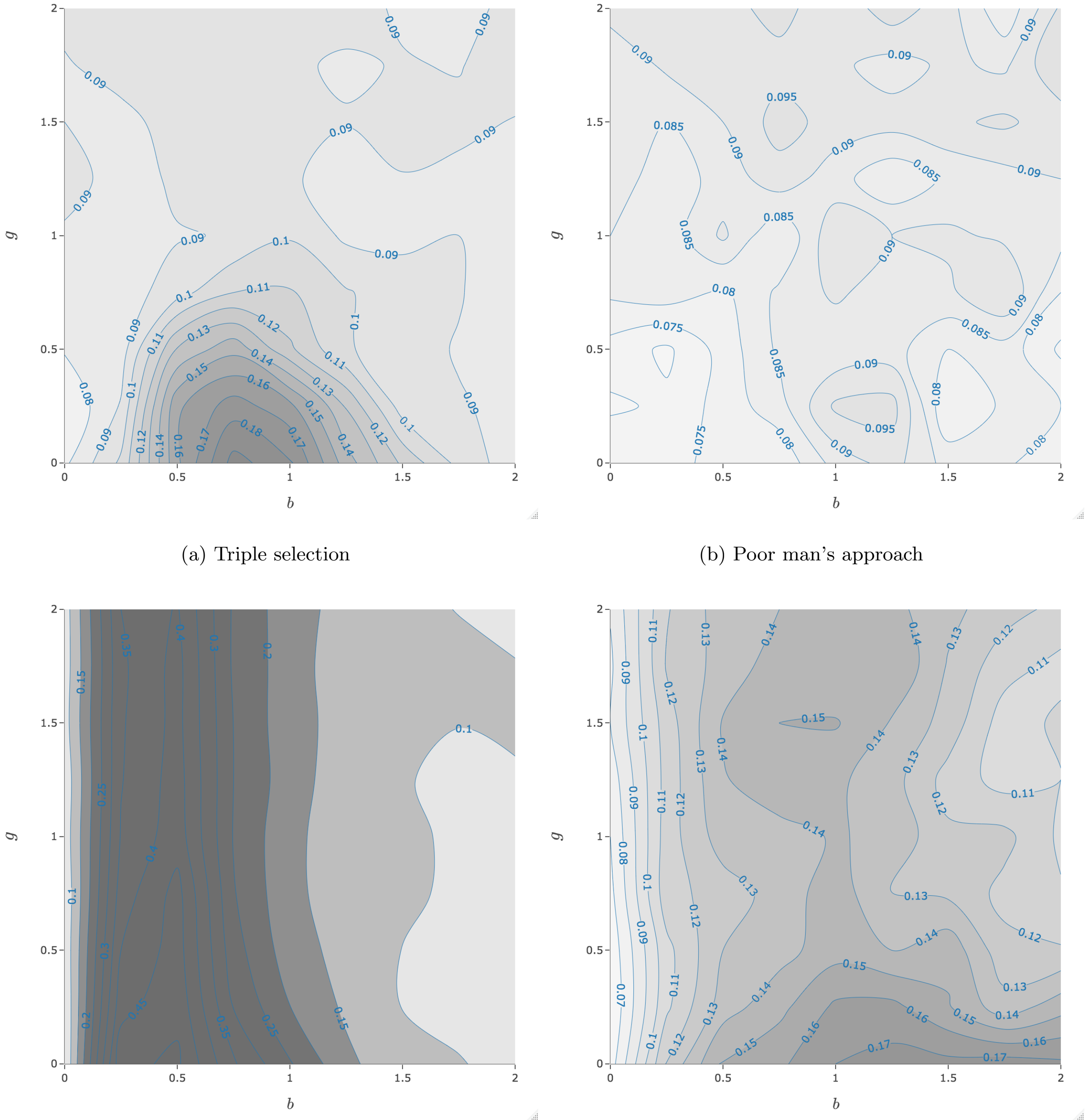


(c) Post-Lasso



(d) Decorrelated test in Fang et al. (2017)

Figure 4: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(a) with $n = 400$, $p = 30$, $\rho = 0.25$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



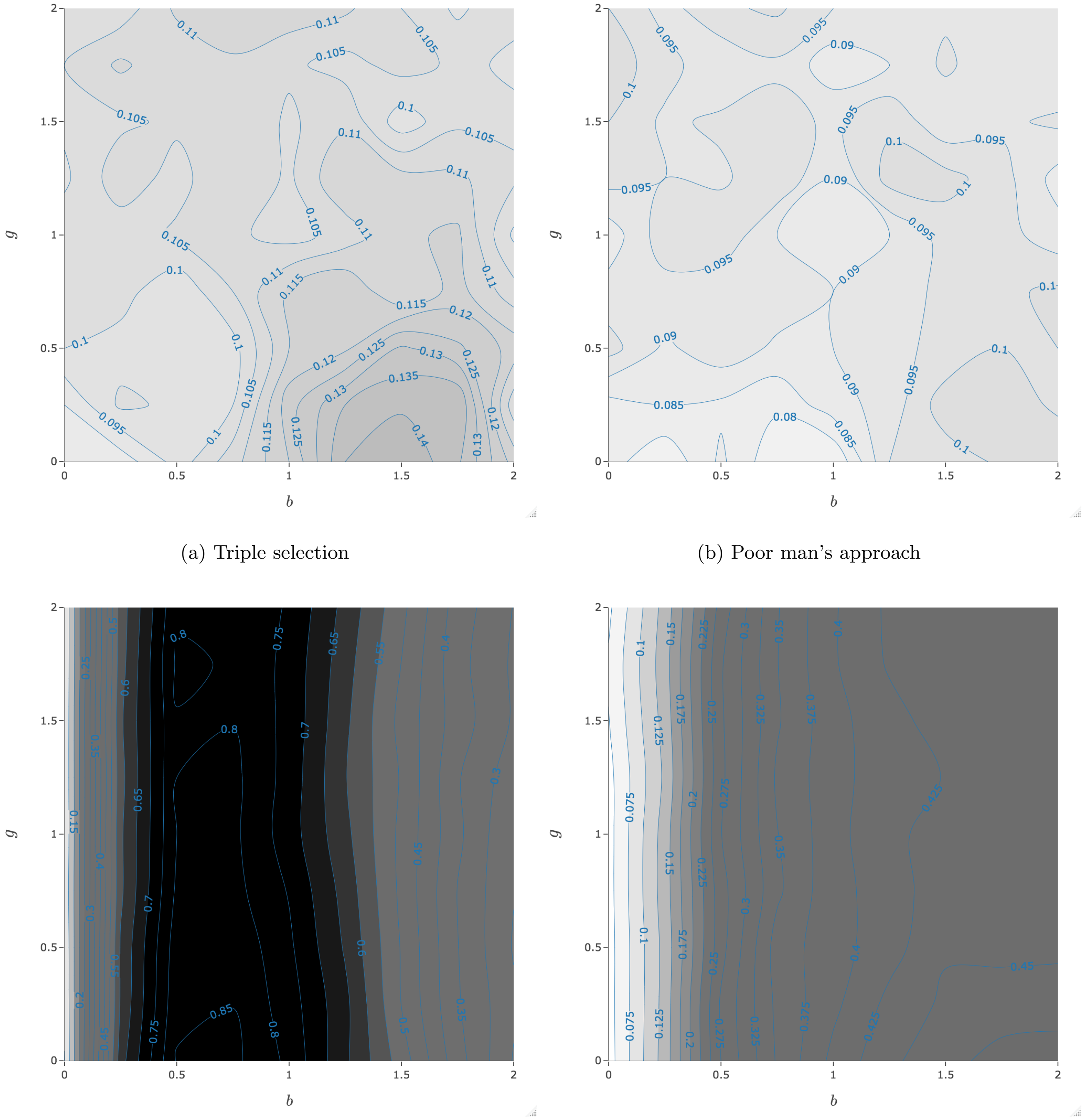
(a) Triple selection

(b) Poor man's approach

(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

Figure 5: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(b) with $n = 400$, $p = 30$, $c_A = 1$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



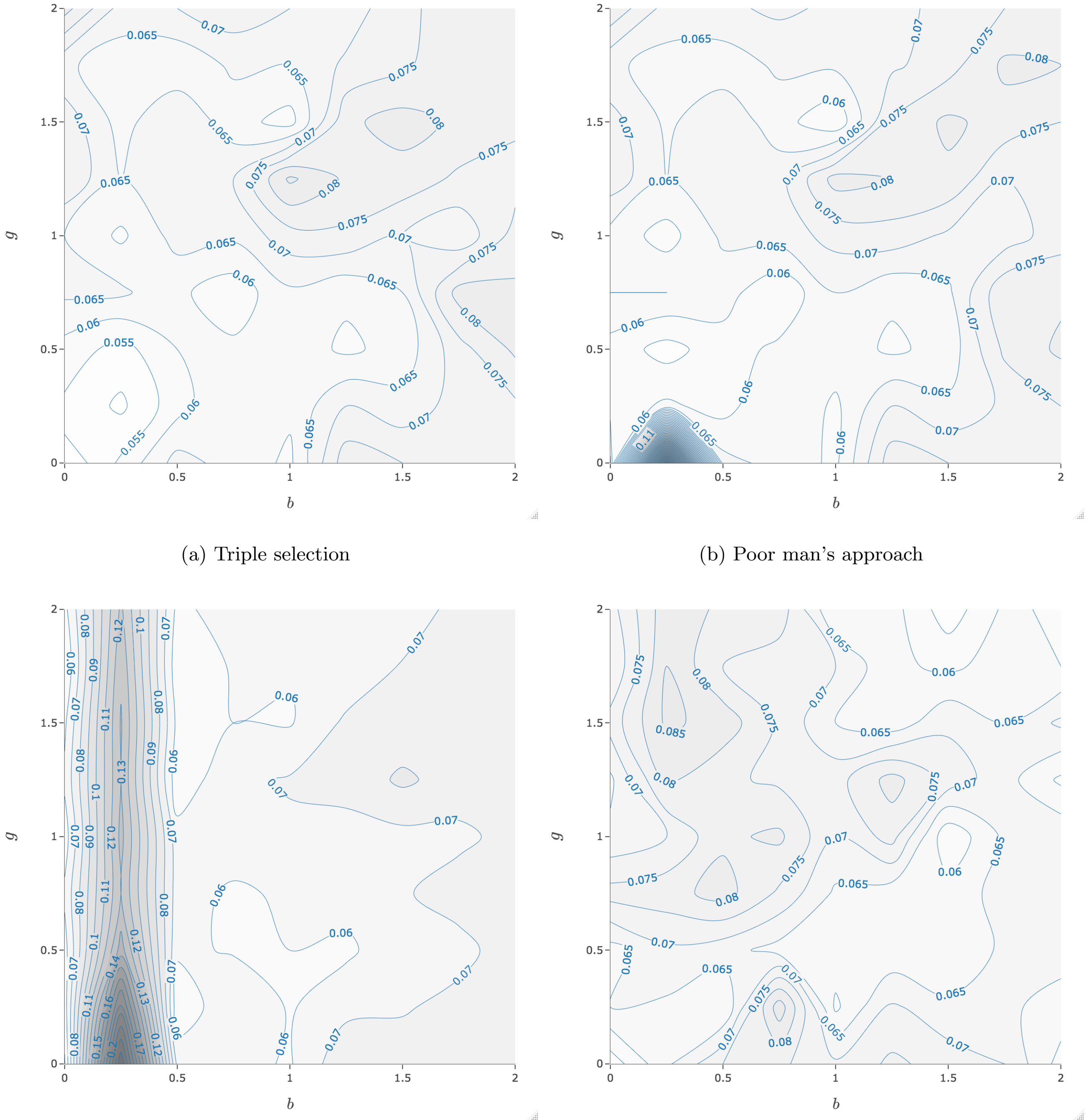
(a) Triple selection

(b) Poor man's approach

(c) Post-Lasso

(d) Decorrelated test inFang et al. (2017)

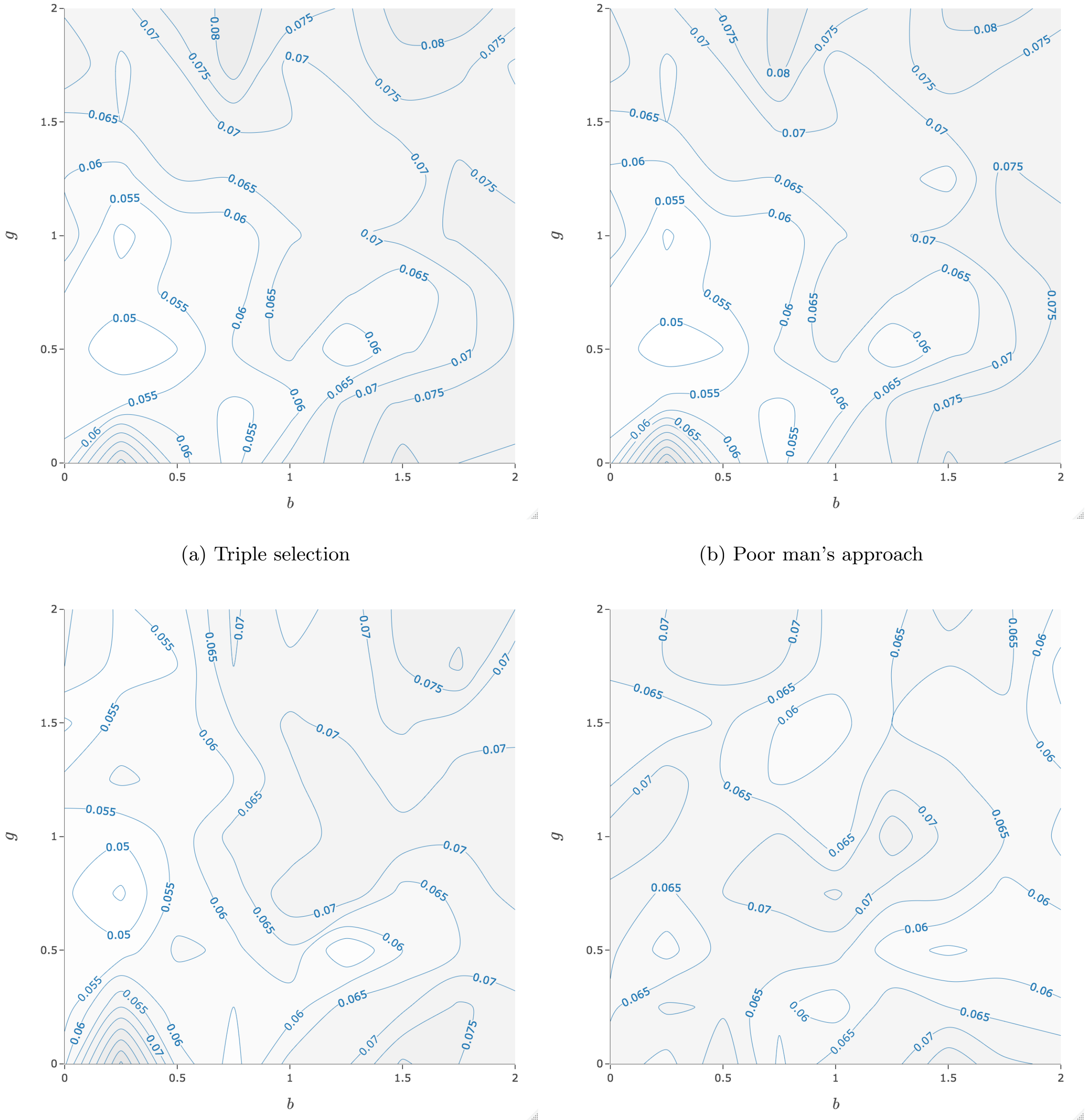
Figure 6: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(b) with $n = 400$, $p = 30$, $c_A = 2$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



(a) Triple selection (b) Poor man's approach

(c) Post-Lasso (d) Decorrelated test in Fang et al. (2017)

Figure 7: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(a) with $n = 400$, $p = 30$, $\rho = 0.50$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



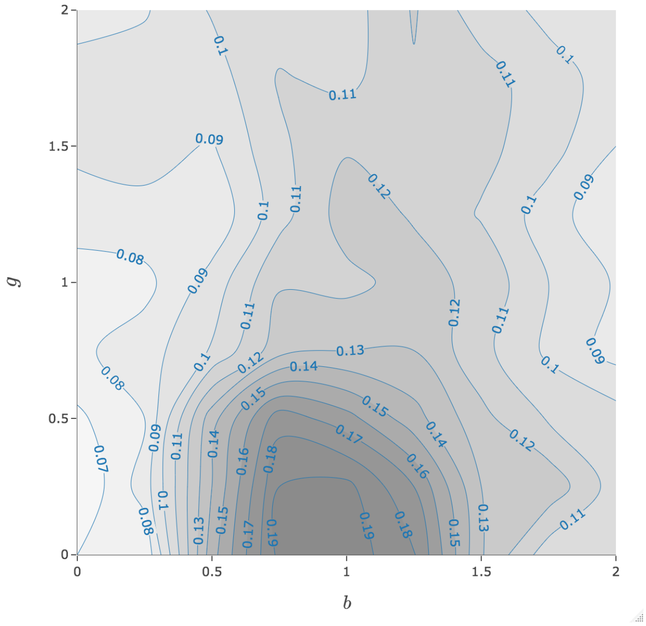
(a) Triple selection

(b) Poor man's approach

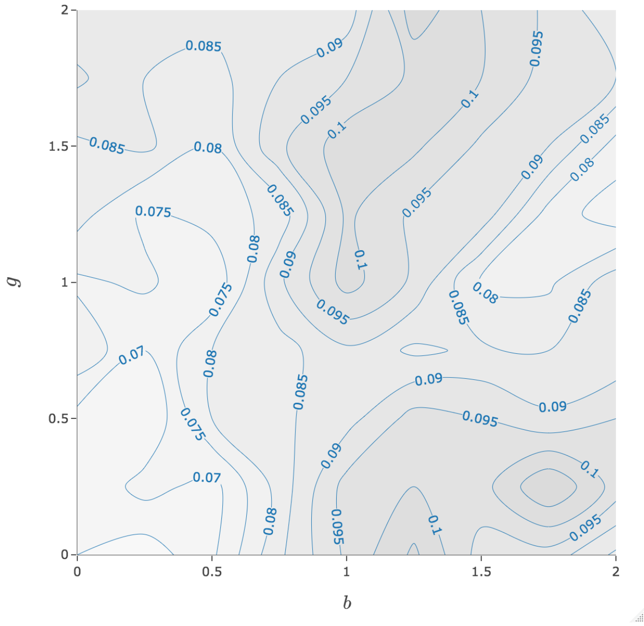
(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

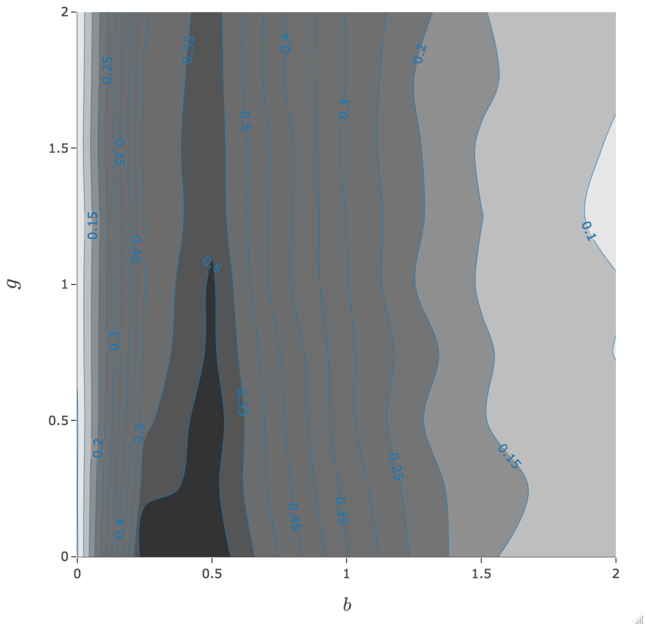
Figure 8: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(a) with $n = 400$, $p = 30$, $\rho = 0.25$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



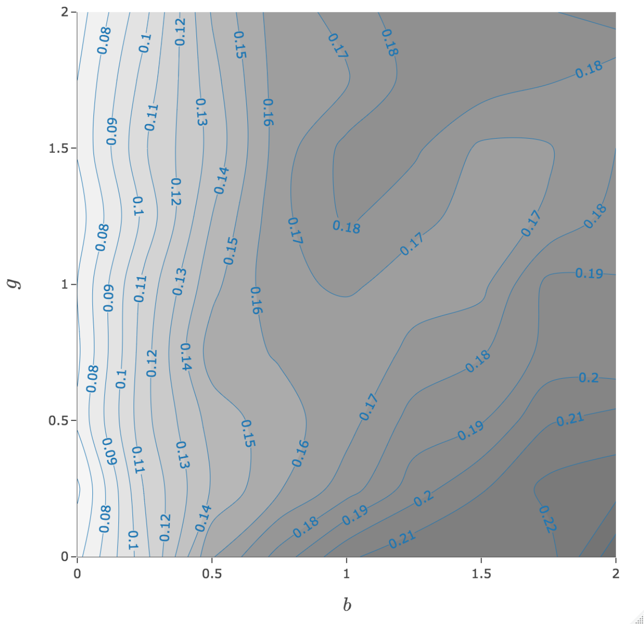
(a) Triple selection



(b) Poor man's approach

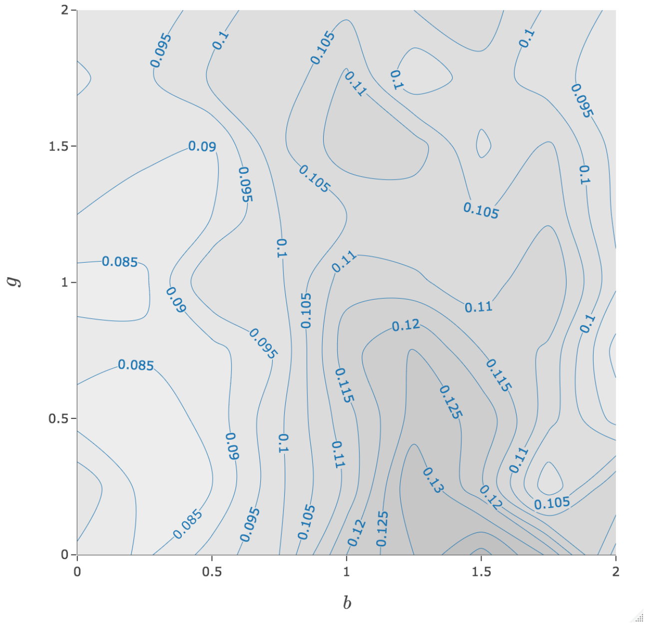


(c) Post-Lasso

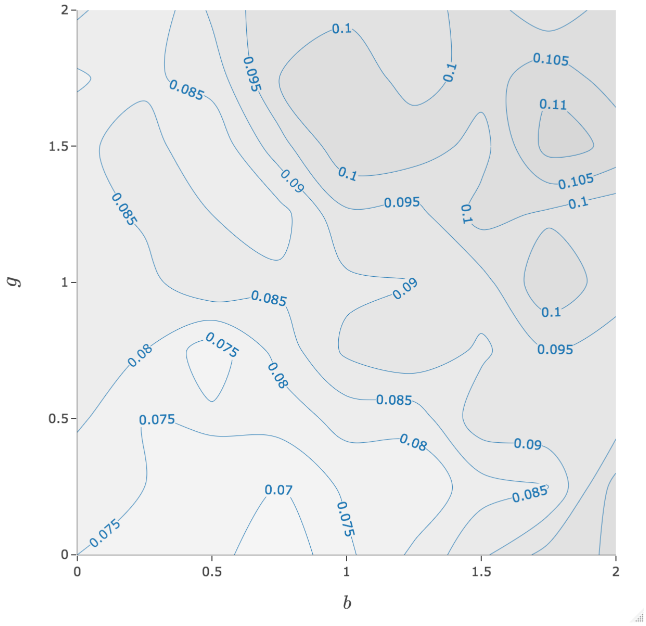


(d) Decorrelated test in Fang et al. (2017)

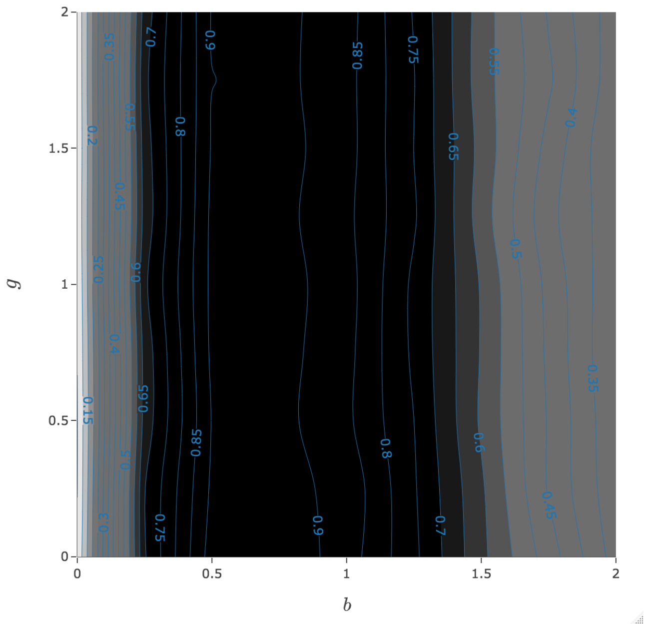
Figure 9: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(b) with $n = 400$, $p = 30$, $c_A = 1$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



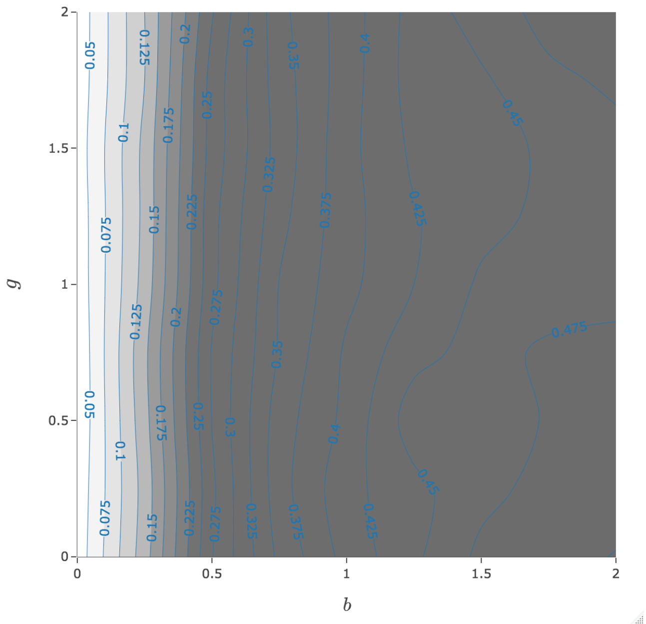
(a) Triple selection



(b) Poor man's approach

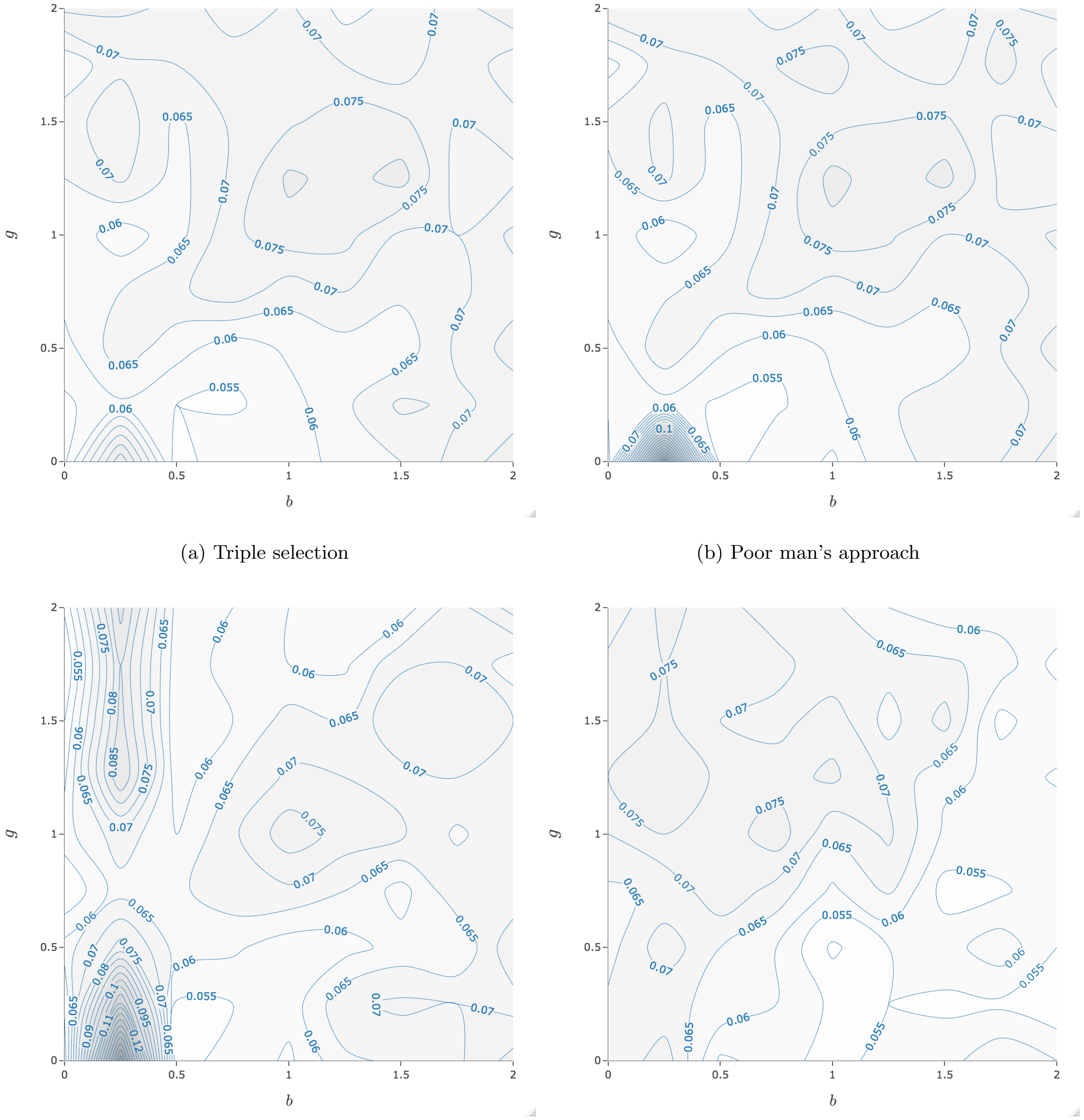


(c) Post-Lasso



(d) Decorrelated test in Fang et al. (2017)

Figure 10: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(b) with $n = 400$, $p = 30$, $c_A = 2$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



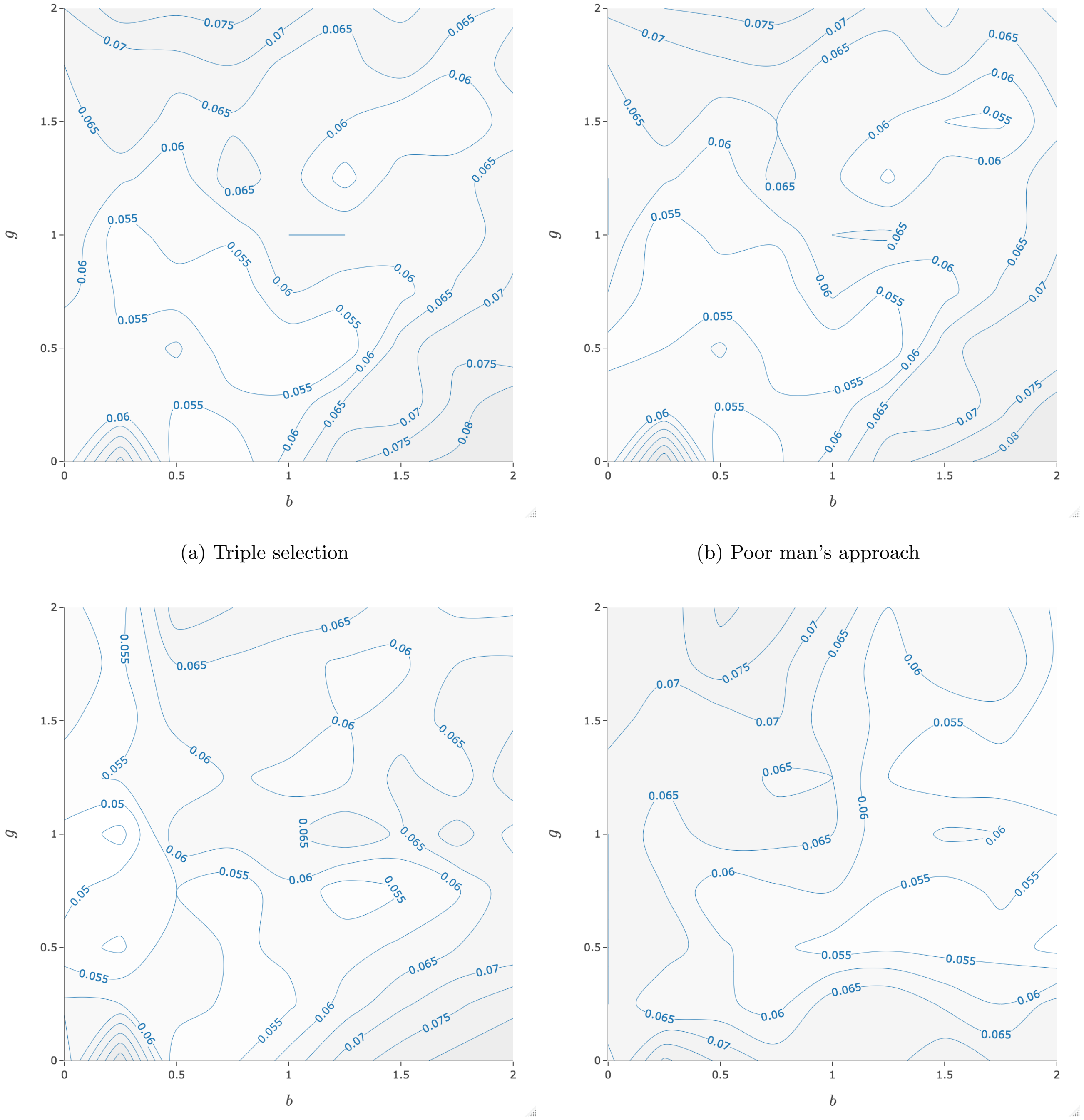
(a) Triple selection

(b) Poor man's approach

(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

Figure 11: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(a) with $n = 400$, $p = 30$, $\rho = 0.50$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



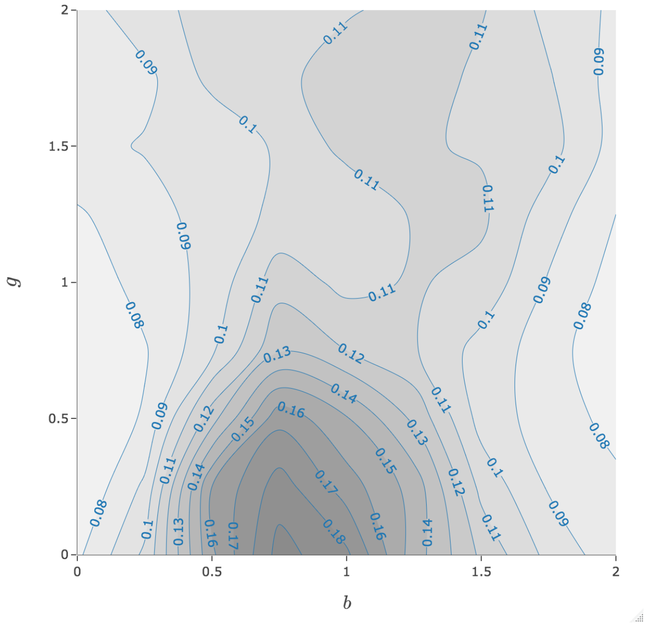
(a) Triple selection

(b) Poor man's approach

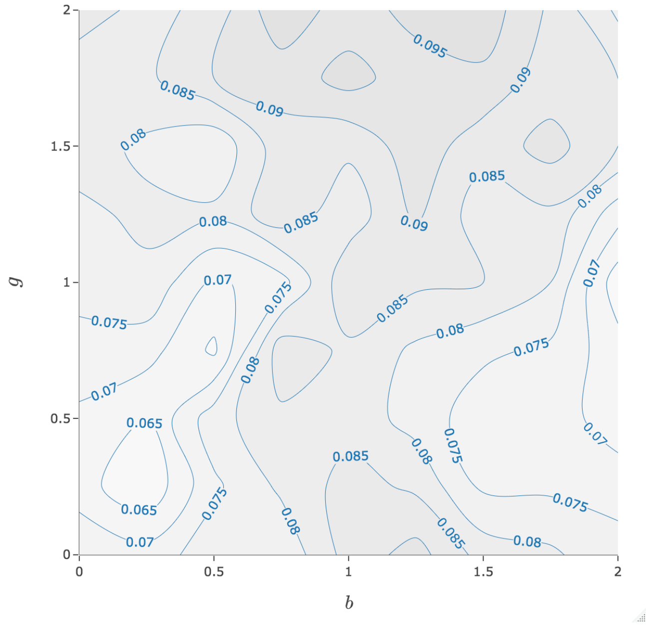
(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

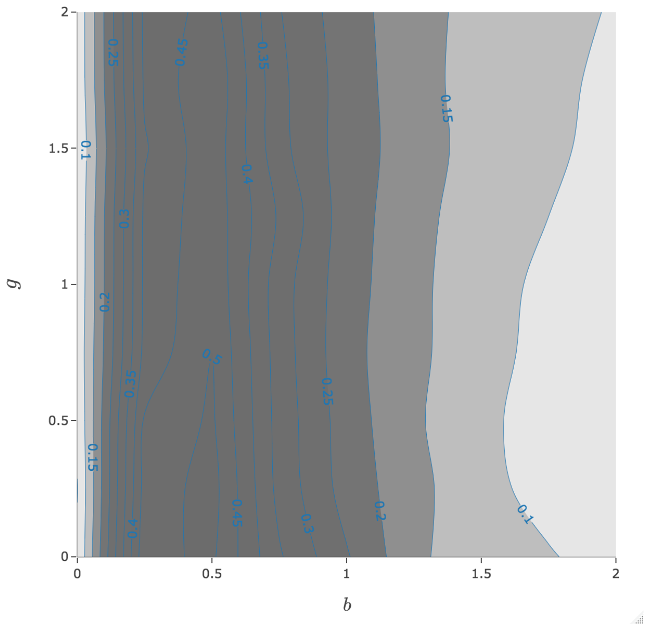
Figure 12: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(a) with $n = 400$, $p = 30$, $\rho = 0.25$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



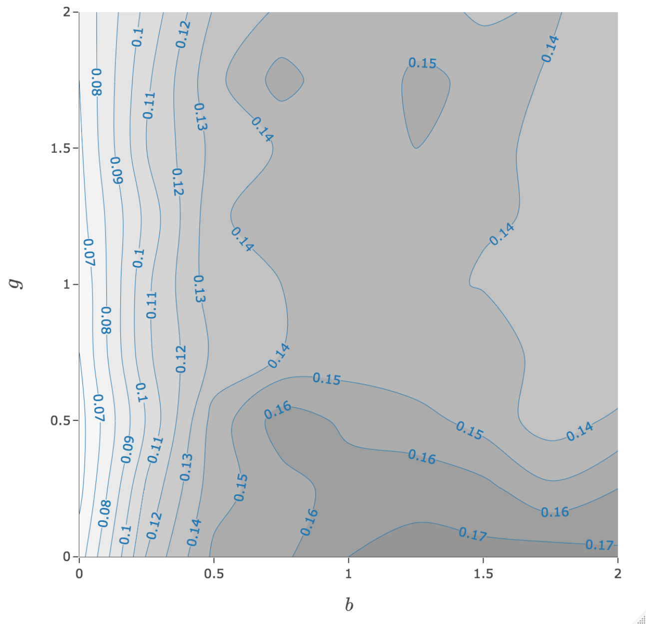
(a) Triple selection



(b) Poor man's approach

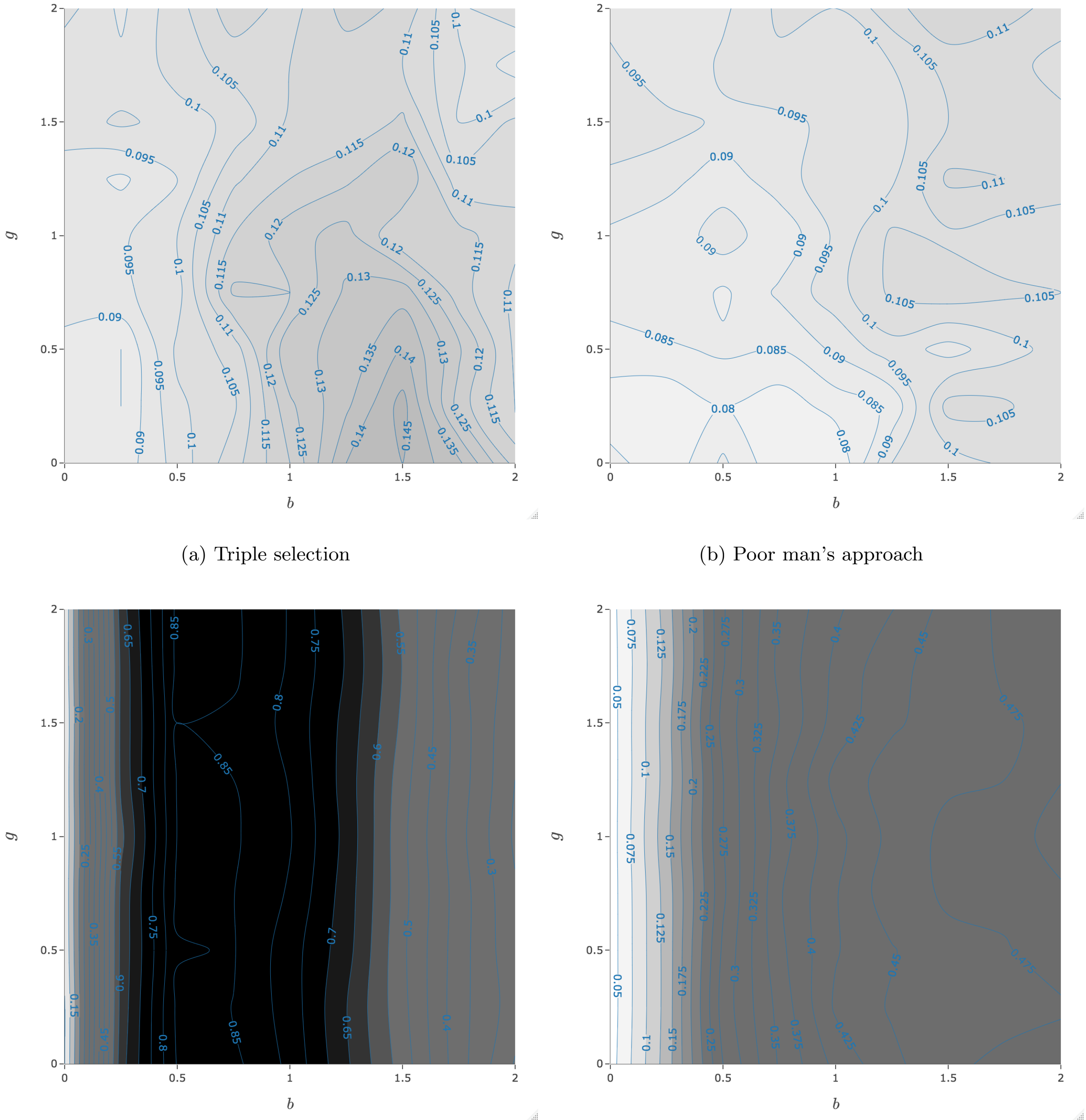


(c) Post-Lasso



(d) Decorrelated test in Fang et al. (2017)

Figure 13: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(b) with $n = 400$, $p = 30$, $c_A = 1$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



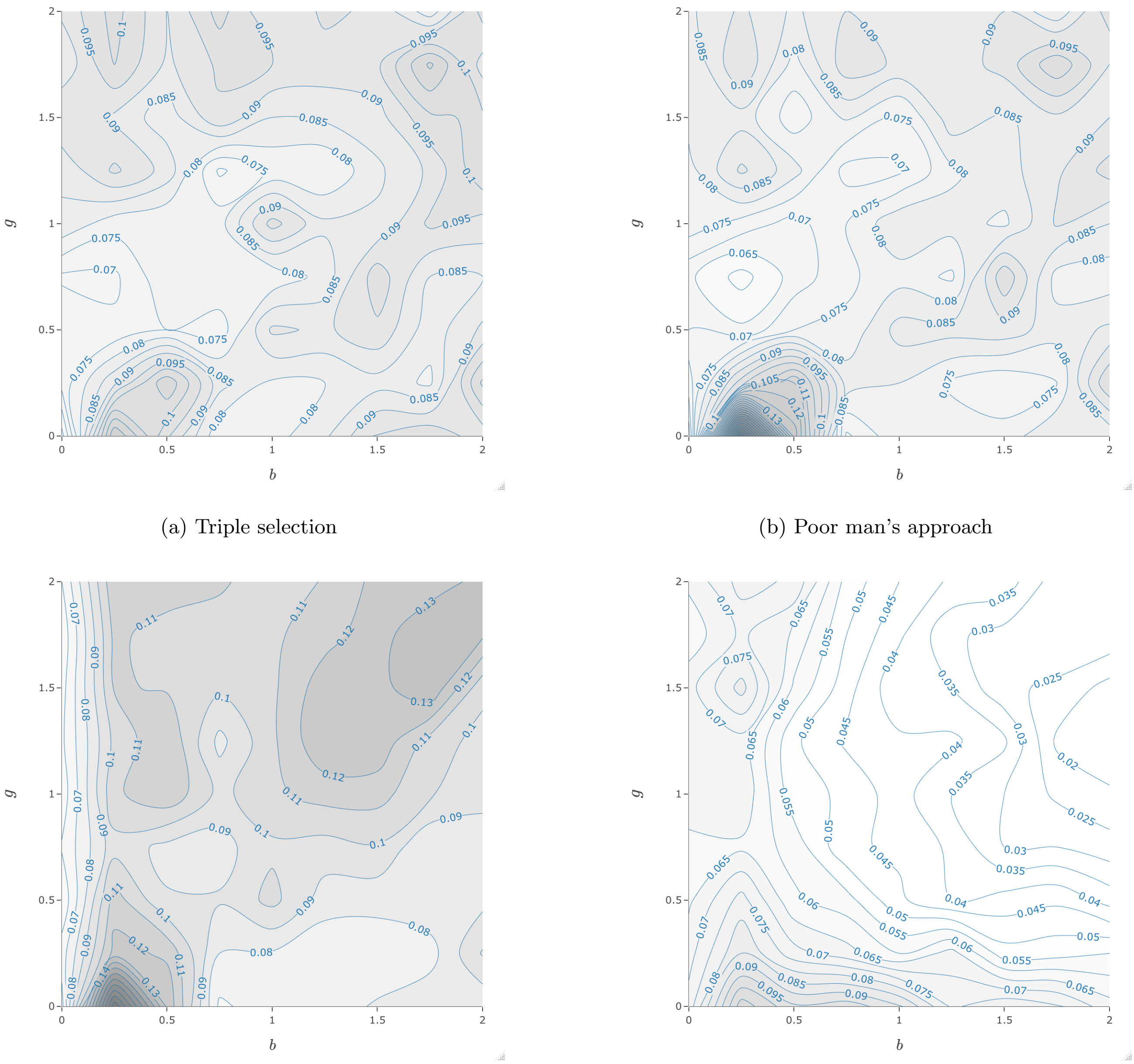
(a) Triple selection

(b) Poor man's approach

(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

Figure 14: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(b) with $n = 400$, $p = 30$, $c_A = 2$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



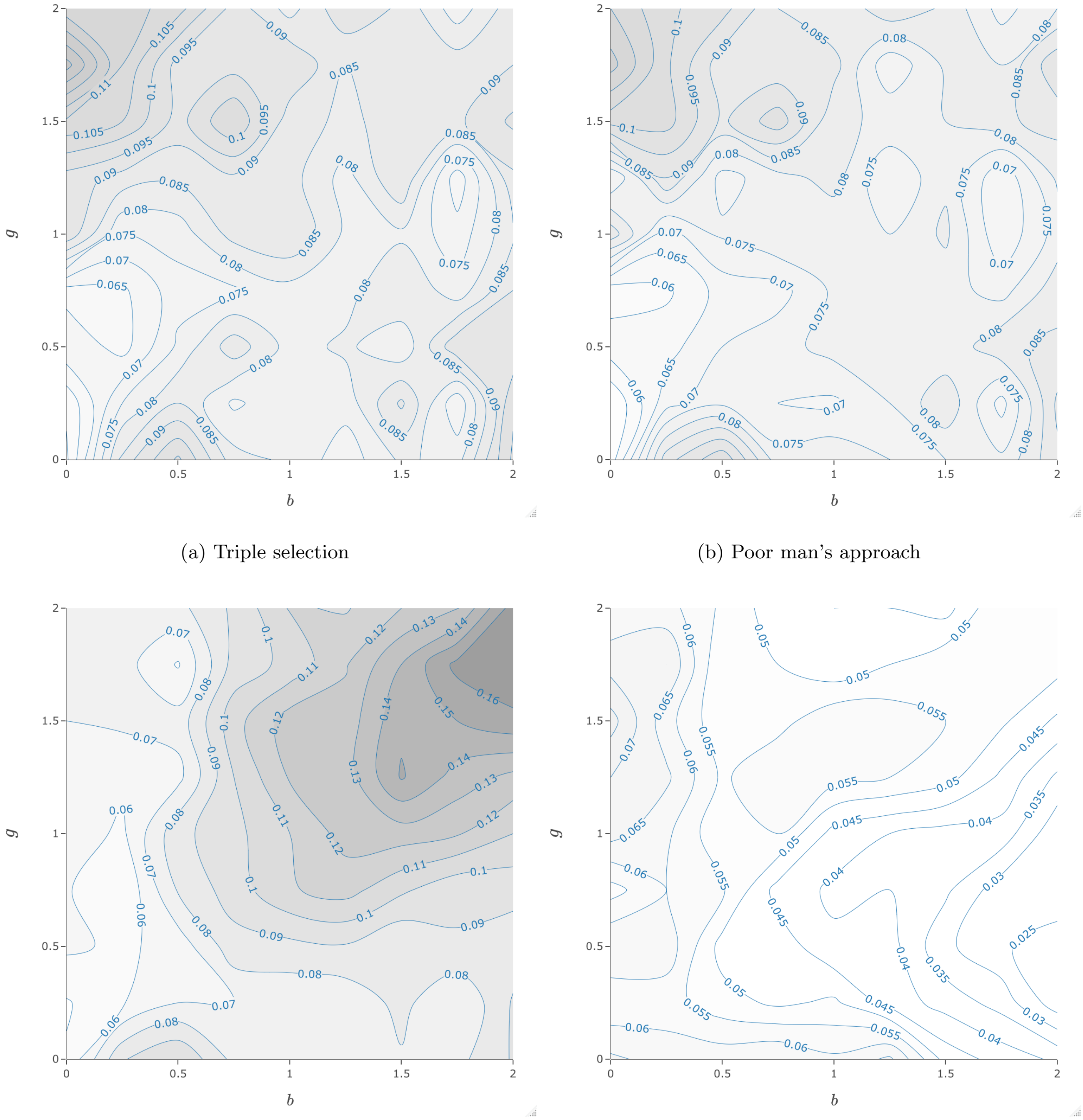
(a) Triple selection

(b) Poor man's approach

(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

Figure 15: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(a) with $n = 200$, $p = 250$, $\rho = 0.50$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



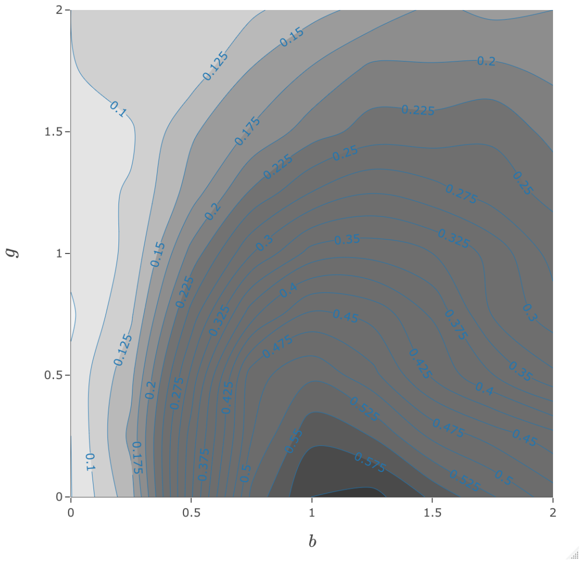
(a) Triple selection

(b) Poor man's approach

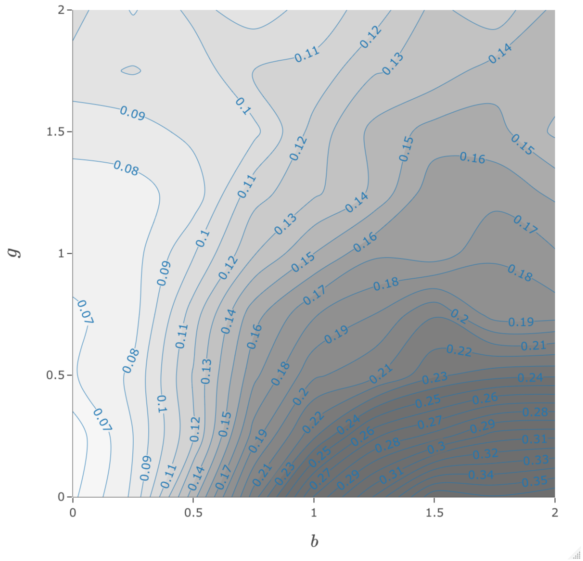
(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

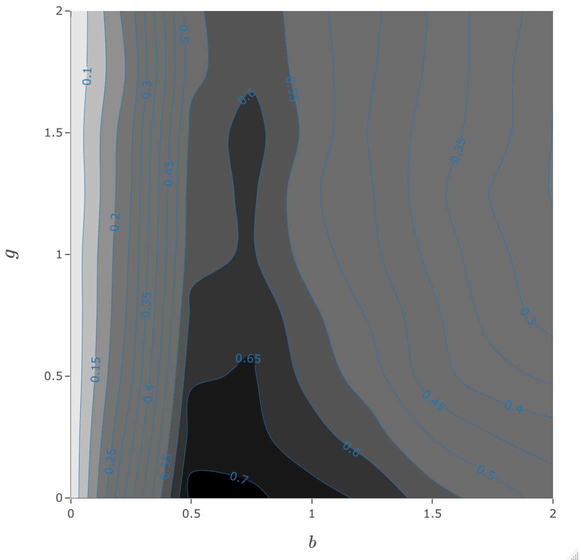
Figure 16: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(a) with $n = 200$, $p = 250$, $\rho = 0.25$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



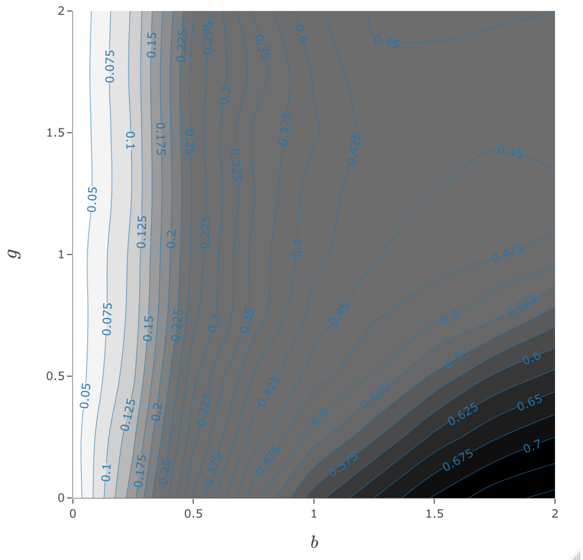
(a) Triple selection



(b) Poor man's approach

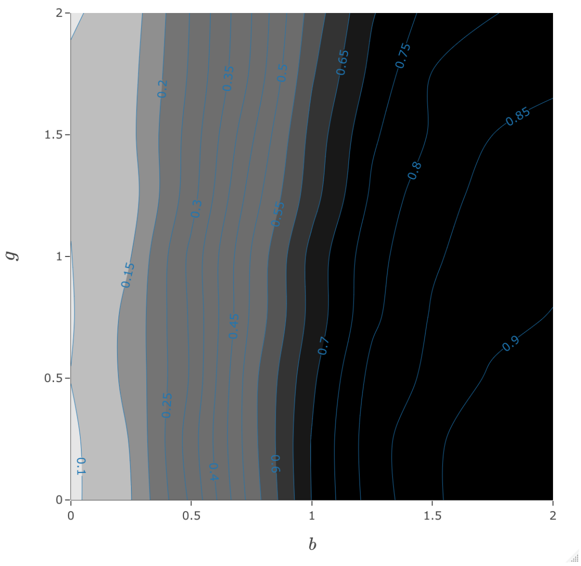


(c) Post-Lasso

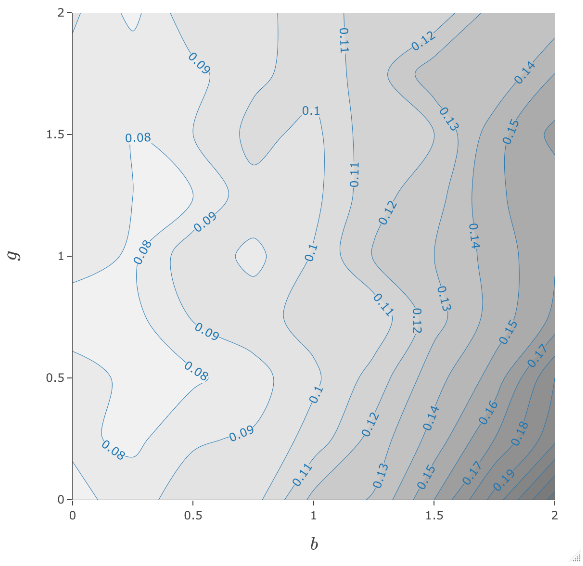


(d) Decorrelated test in Fang et al. (2017)

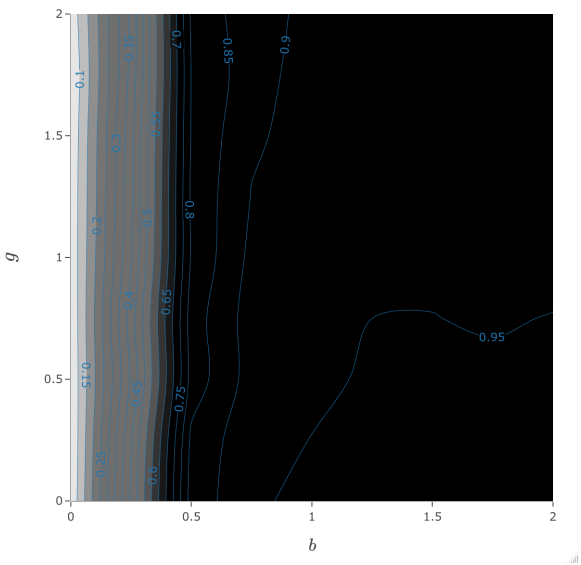
Figure 17: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(b) with $n = 200$, $p = 250$, $c_A = 1$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



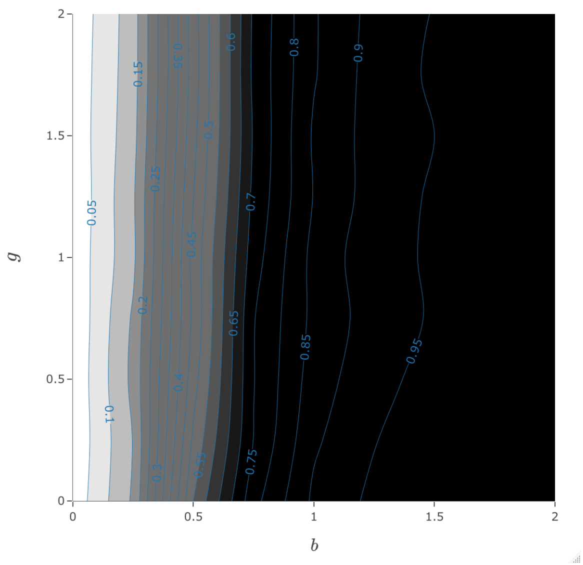
(a) Triple selection



(b) Poor man's approach

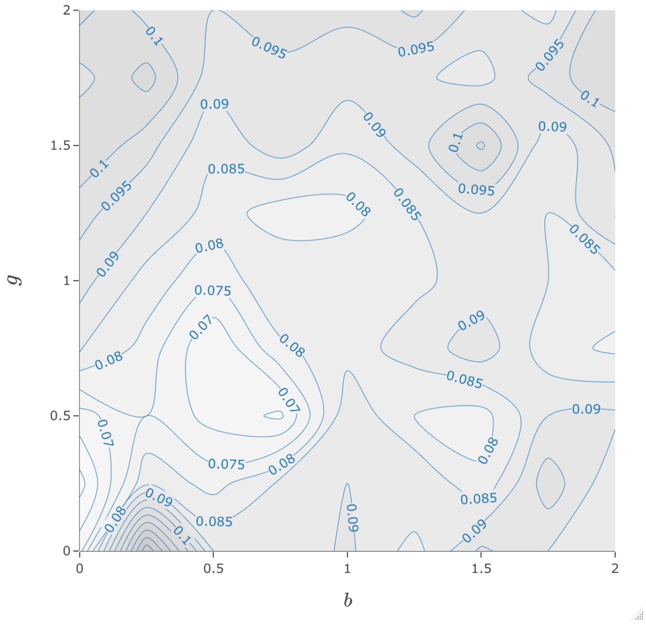


(c) Post-Lasso

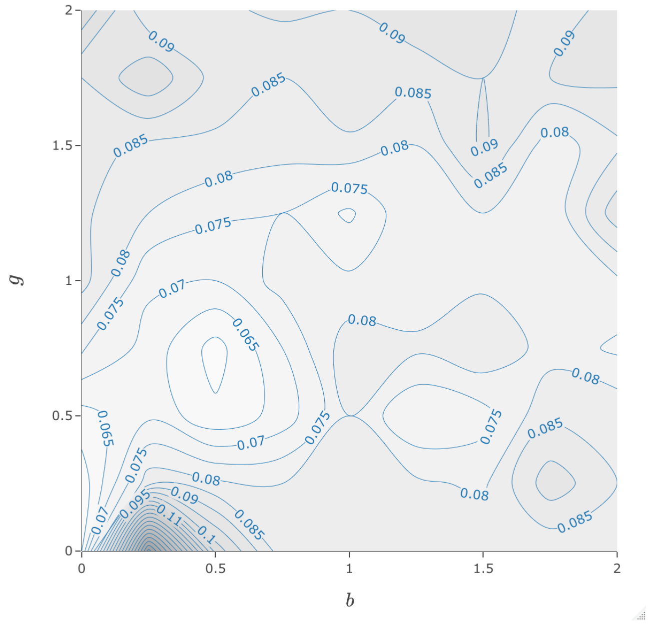


(d) Decorrelated test in Fang et al. (2017)

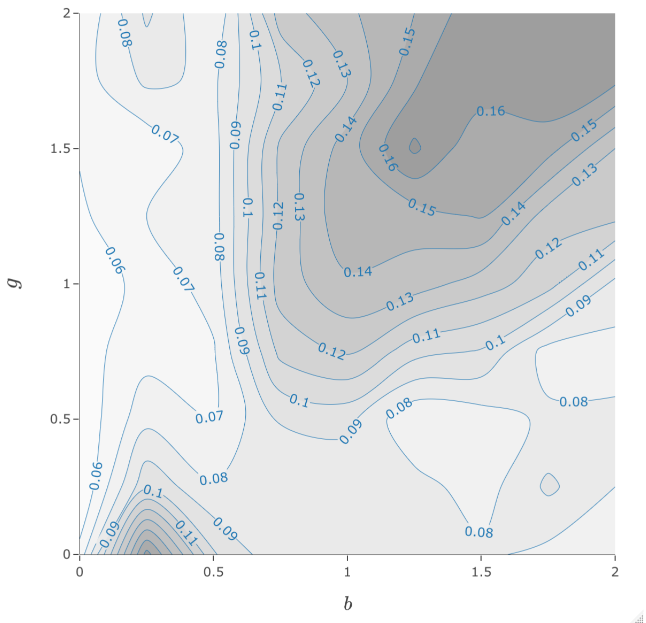
Figure 18: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(b) with $n = 200$, $p = 250$, $c_A = 2$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



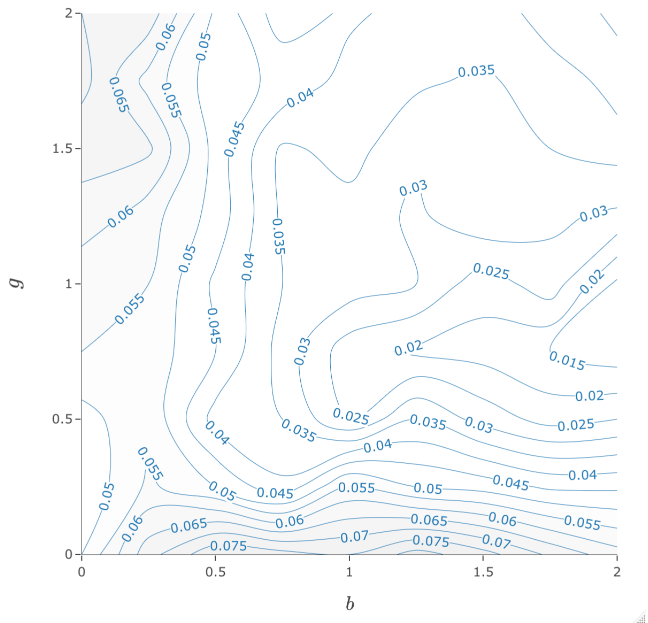
(a) Triple selection



(b) Poor man's approach

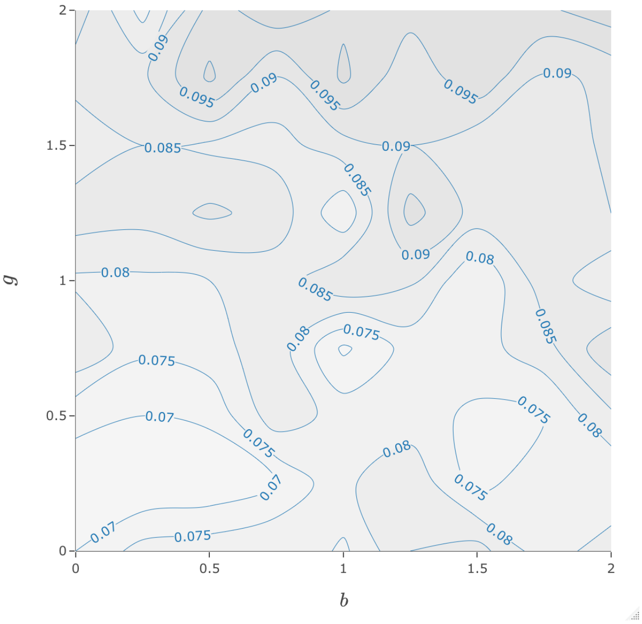


(c) Post-Lasso

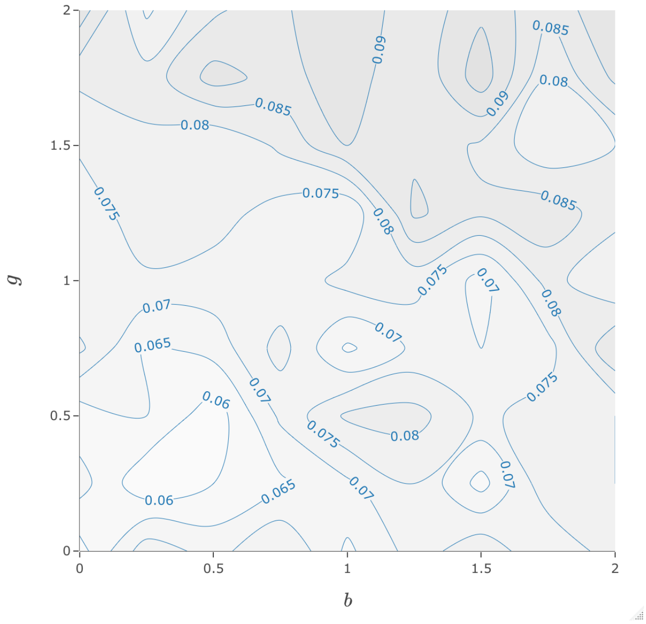


(d) Decorrelated test in Fang et al. (2017)

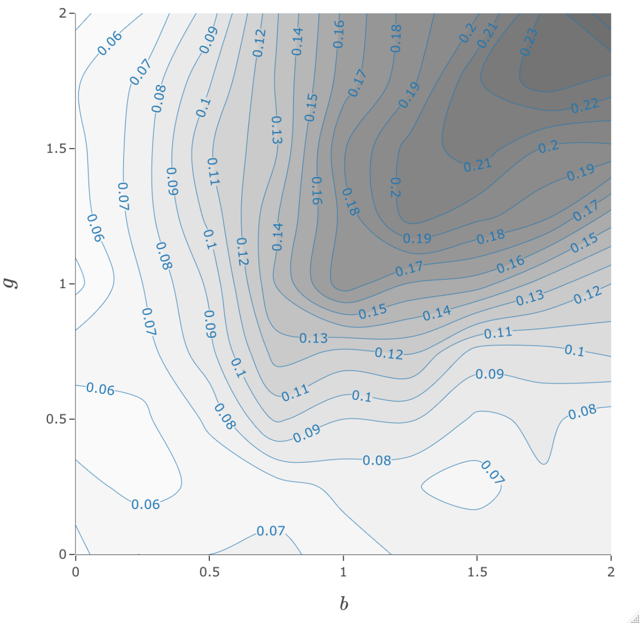
Figure 19: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(a) with $n = 200$, $p = 250$, $\rho = 0.50$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



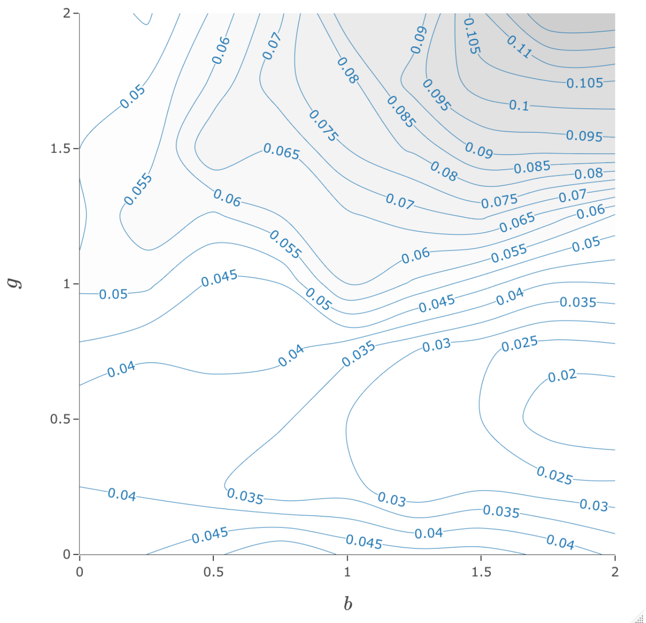
(a) Triple selection



(b) Poor man's approach

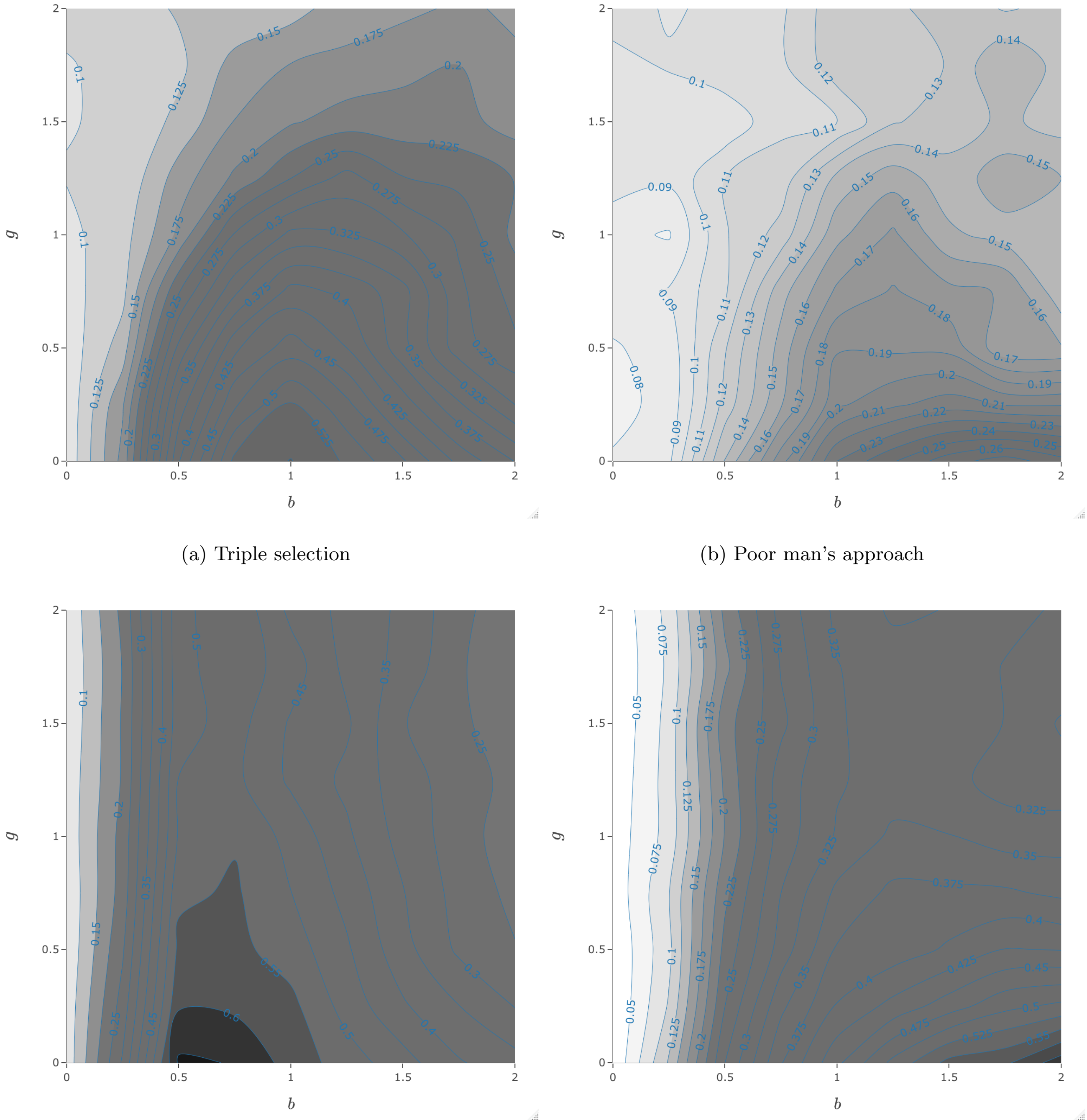


(c) Post-Lasso



(d) Decorrelated test in Fang et al. (2017)

Figure 20: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(a) with $n = 200$, $p = 250$, $\rho = 0.25$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



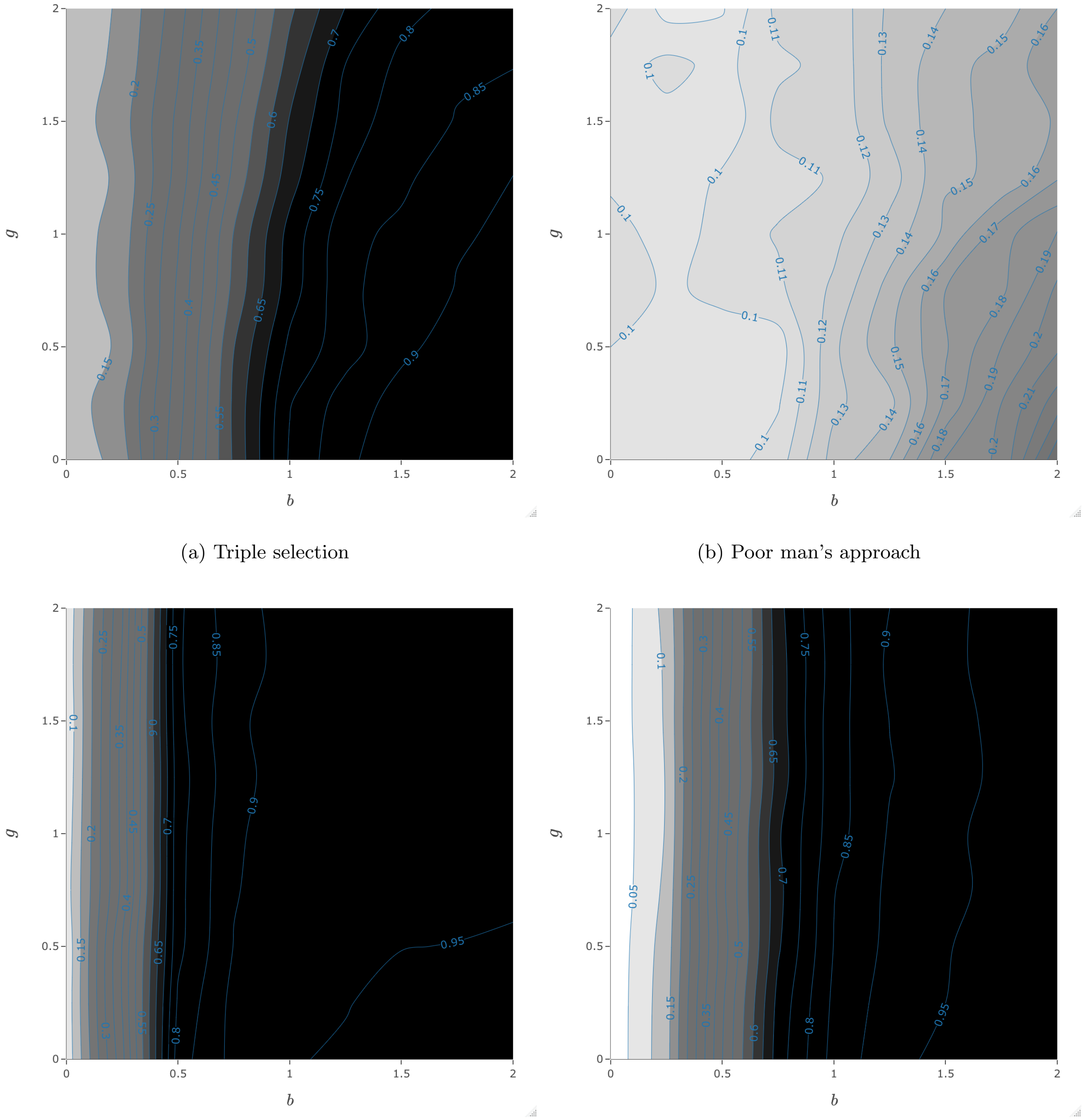
(a) Triple selection

(b) Poor man's approach

(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

Figure 21: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(b) with $n = 200$, $p = 250$, $c_A = 1$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



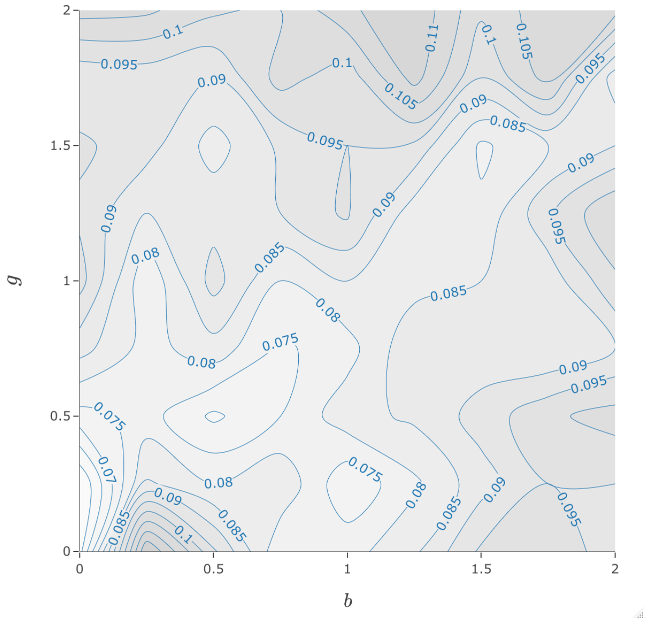
(a) Triple selection

(b) Poor man's approach

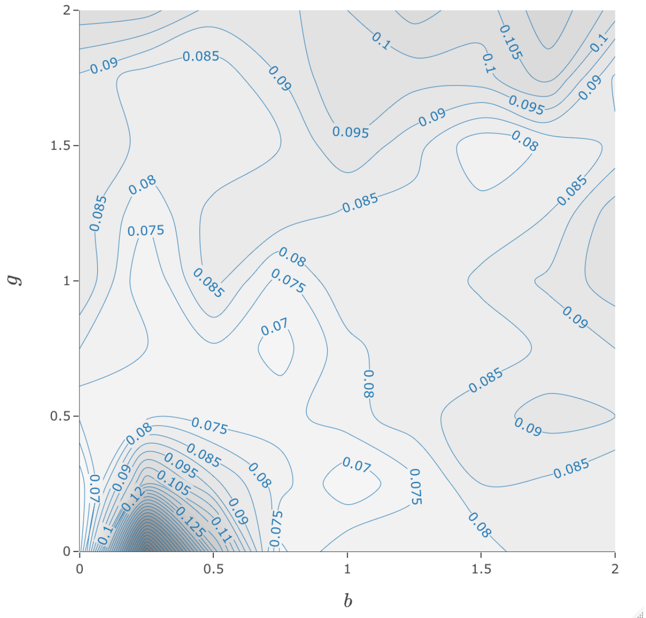
(c) Post-Lasso

(d) Decorrelated test inFang et al. (2017)

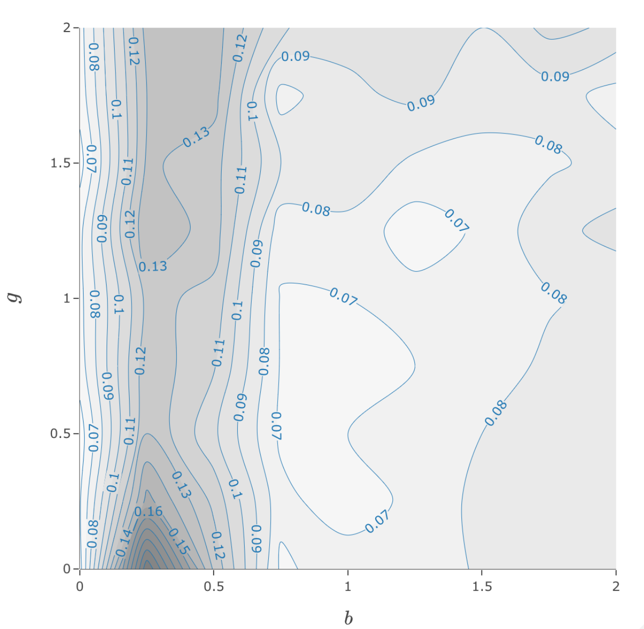
Figure 22: Empirical Type I error rate at the 5% significance level of the different tests under Setting 1(b) with $n = 200$, $p = 250$, $c_A = 2$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



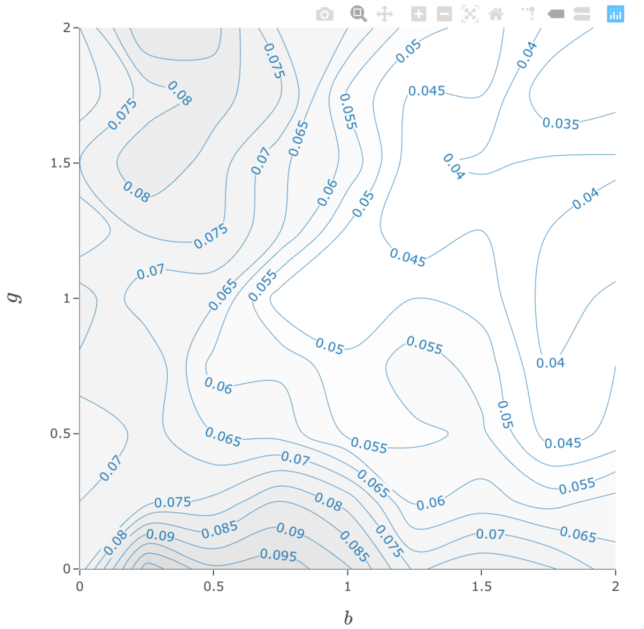
(a) Triple selection



(b) Poor man's approach

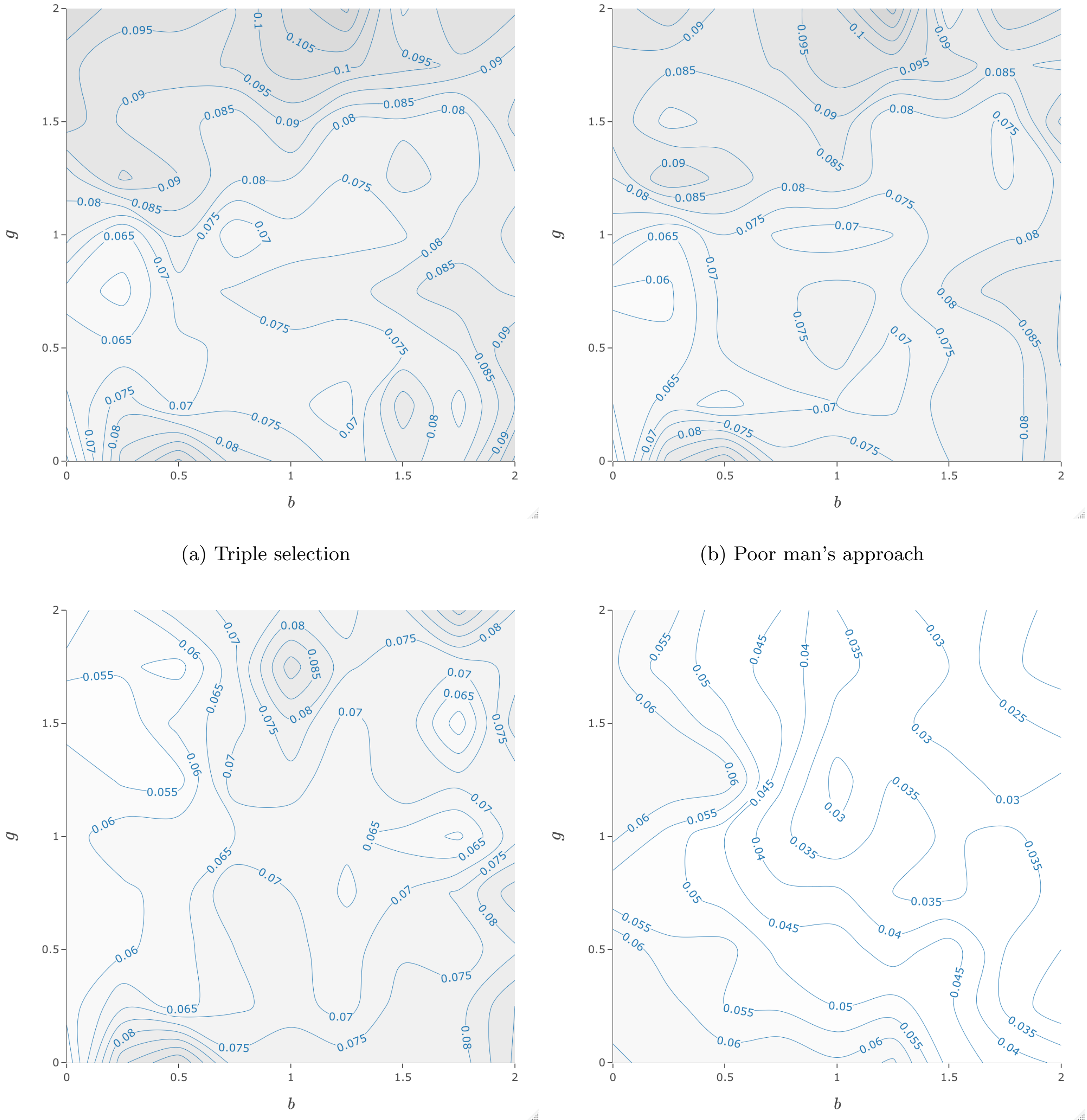


(c) Post-Lasso



(d) Decorrelated test in Fang et al. (2017)

Figure 23: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(a) with $n = 200$, $p = 250$, $\rho = 0.50$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



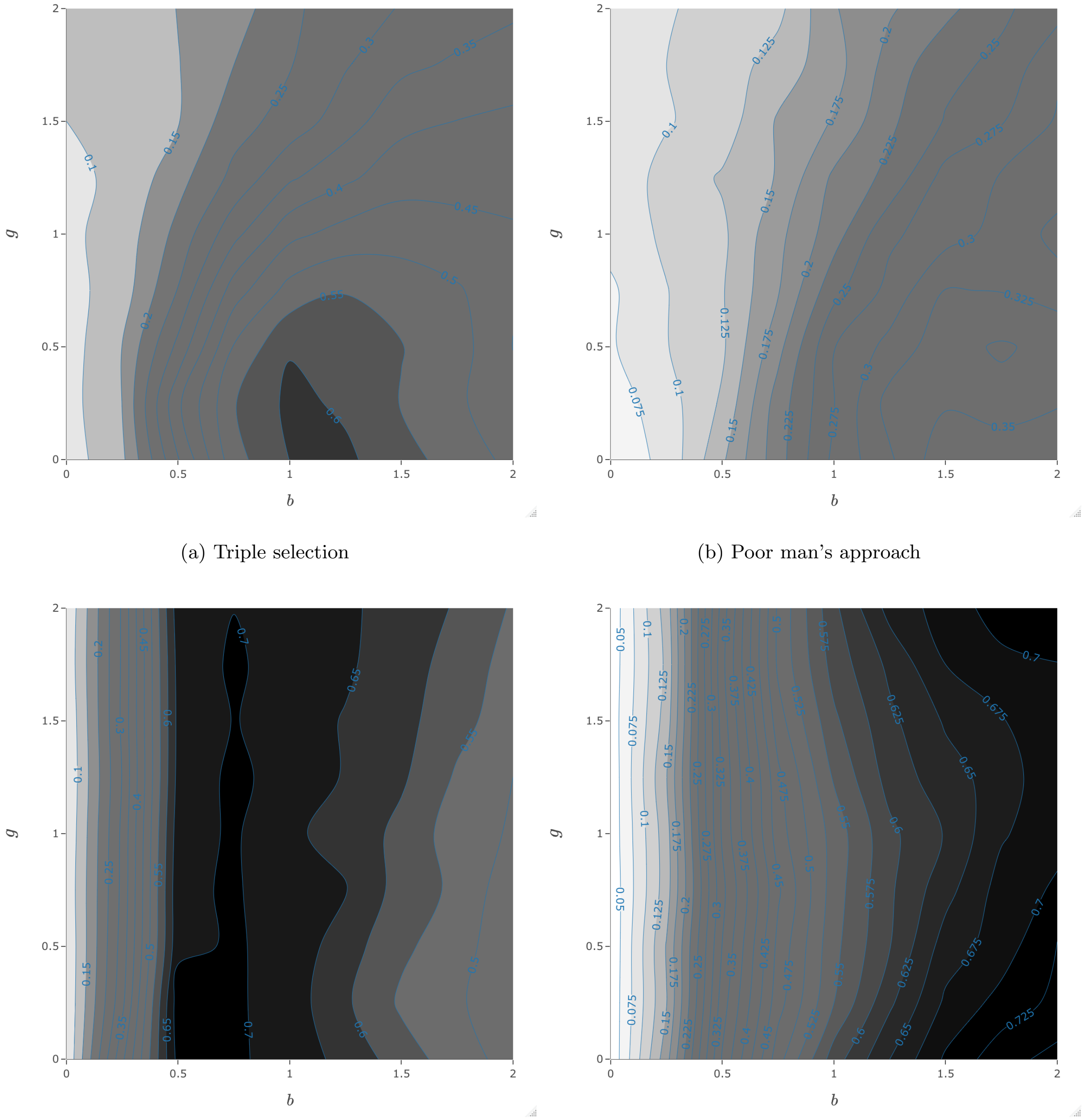
(a) Triple selection

(b) Poor man's approach

(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

Figure 24: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(a) with $n = 200$, $p = 250$, $\rho = 0.25$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



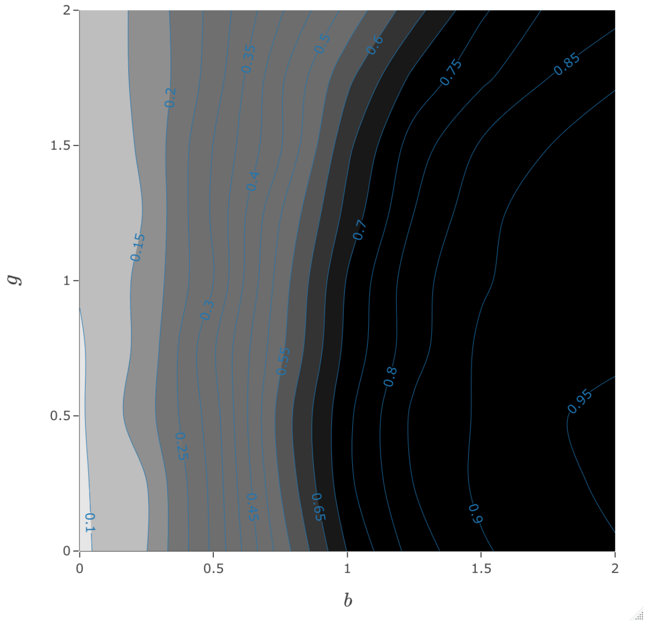
(a) Triple selection

(b) Poor man's approach

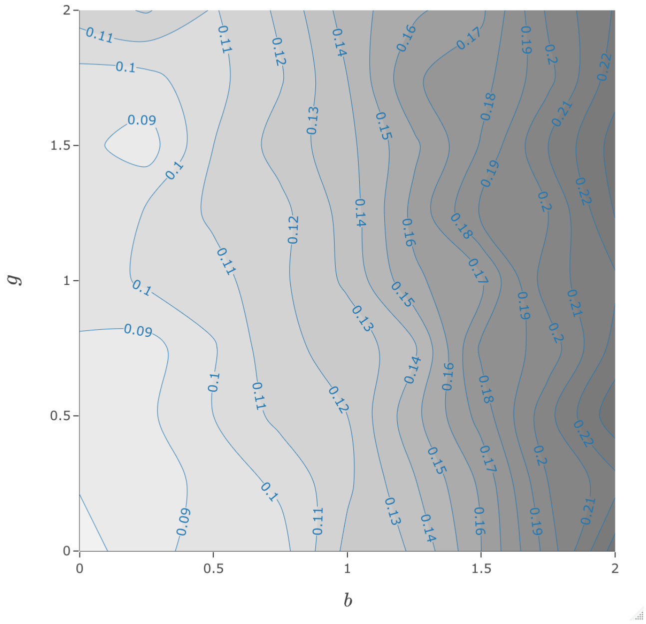
(c) Post-Lasso

(d) Decorrelated test in Fang et al. (2017)

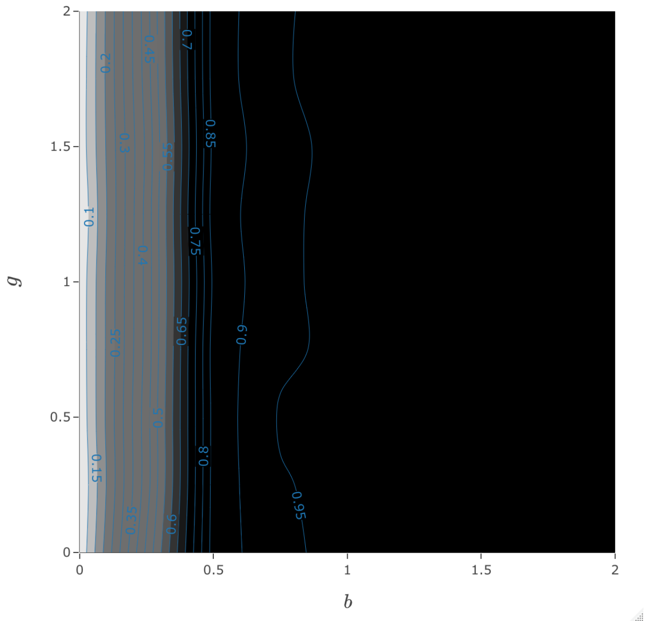
Figure 25: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(b) with $n = 200$, $p = 250$, $c_A = 1$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



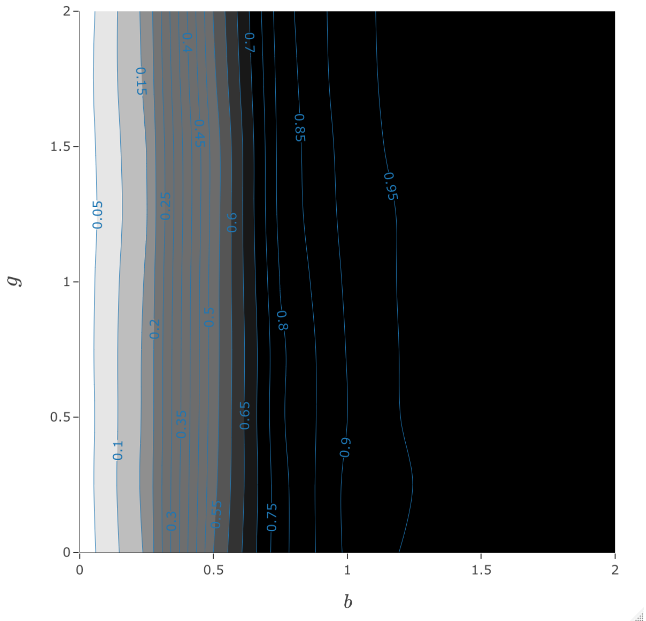
(a) Triple selection



(b) Poor man's approach

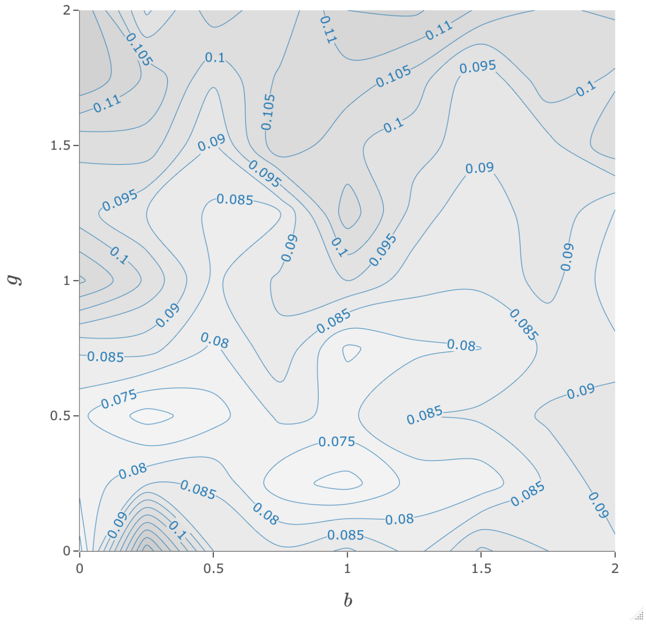


(c) Post-Lasso

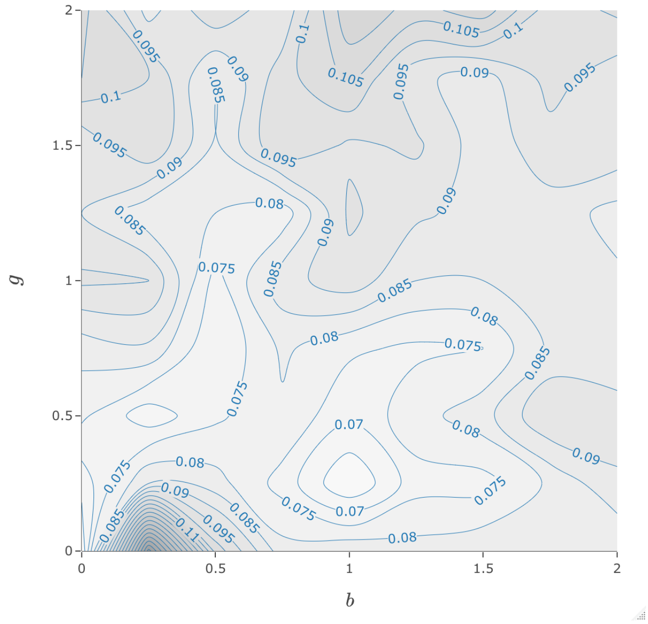


(d) Decorrelated test in Fang et al. (2017)

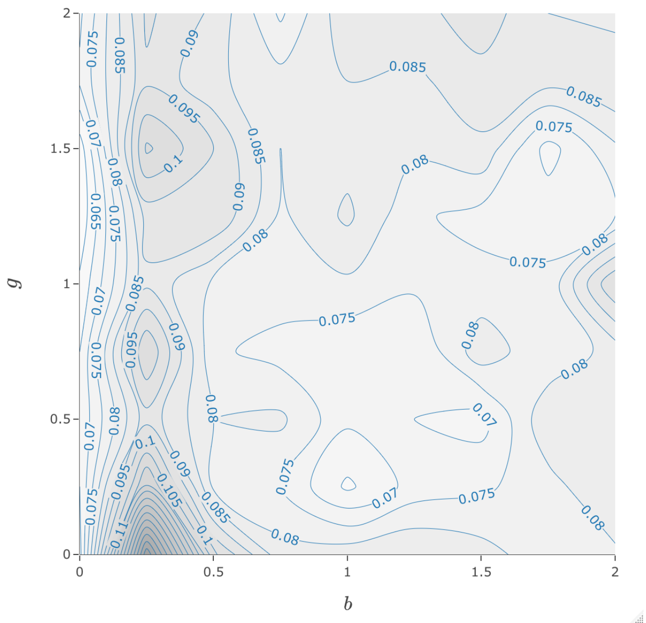
Figure 26: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(b) with $n = 200$, $p = 250$, $c_A = 2$, $\eta_1 = 1$, $\beta_0 = 0$ and $\eta_0 = 0$.



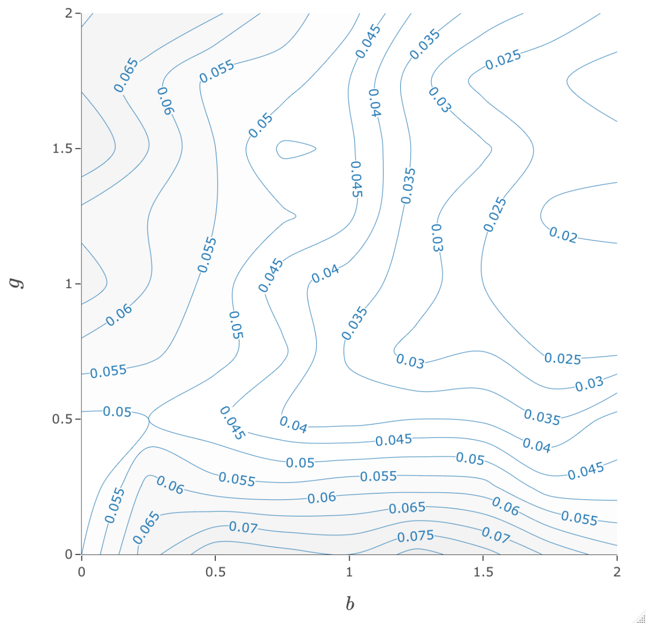
(a) Triple selection



(b) Poor man's approach

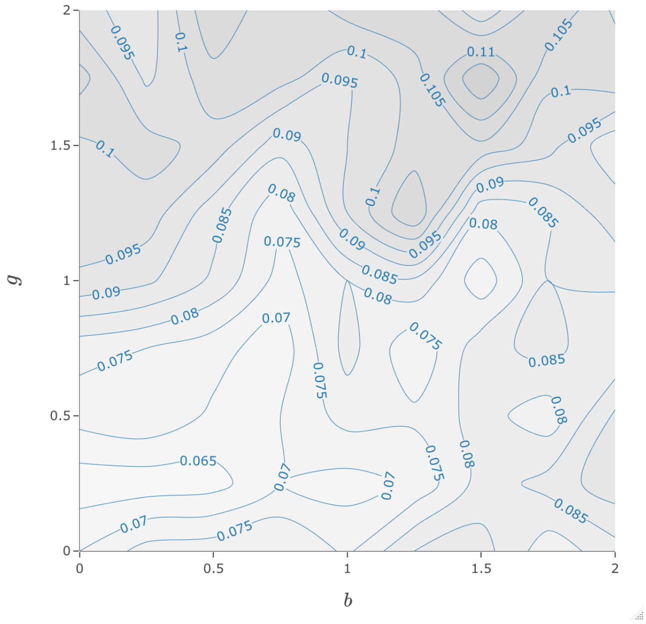


(c) Post-Lasso

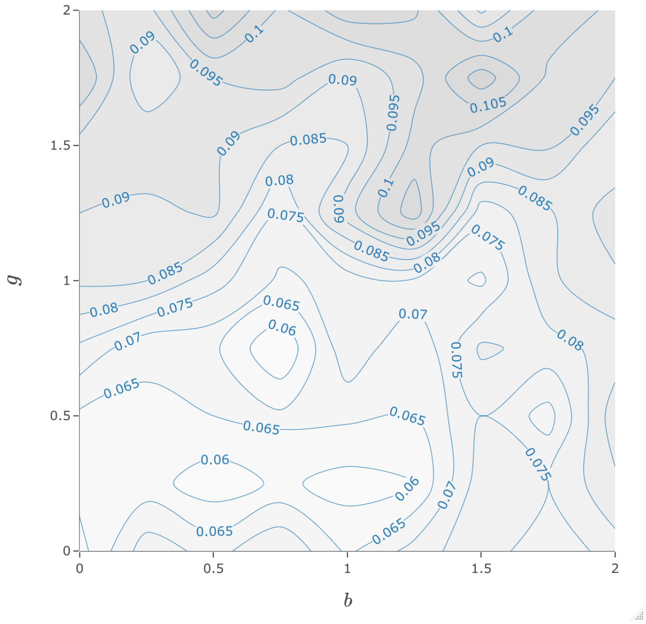


(d) Decorrelated test in Fang et al. (2017)

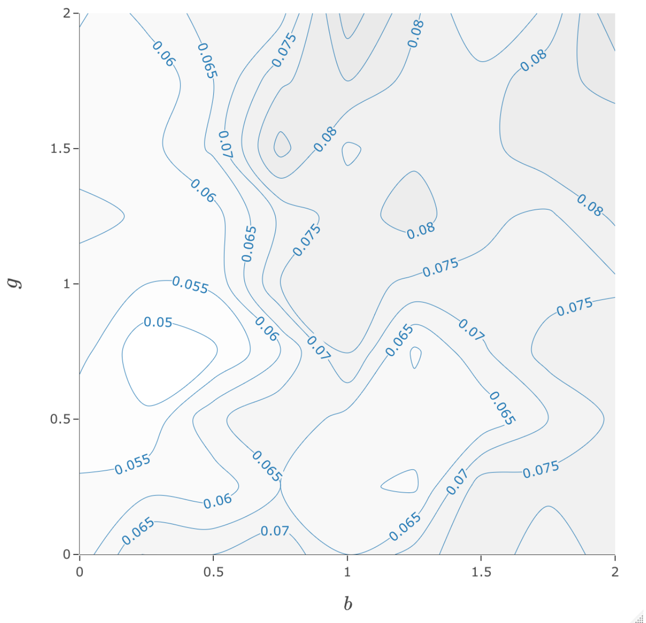
Figure 27: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(a) with $n = 200$, $p = 250$, $\rho = 0.50$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



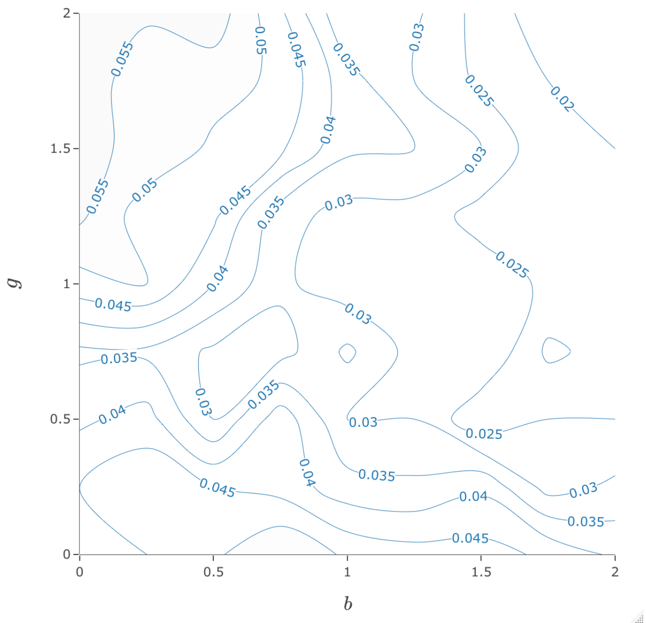
(a) Triple selection



(b) Poor man's approach

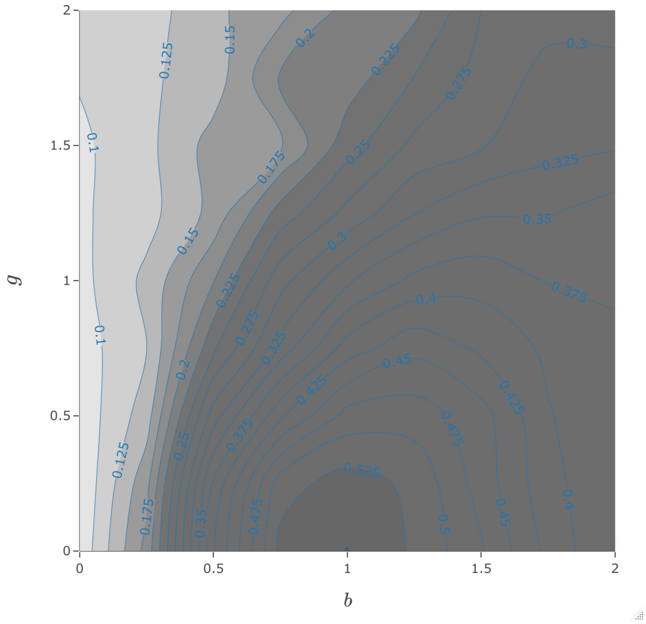


(c) Post-Lasso

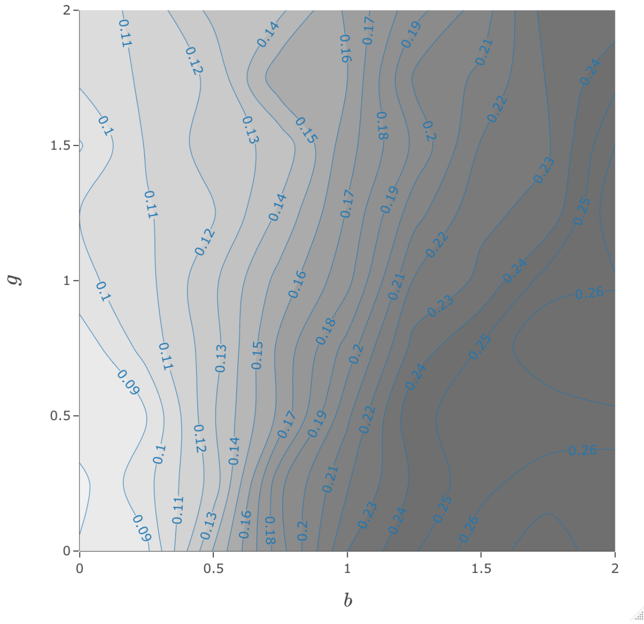


(d) Decorrelated test in Fang et al. (2017)

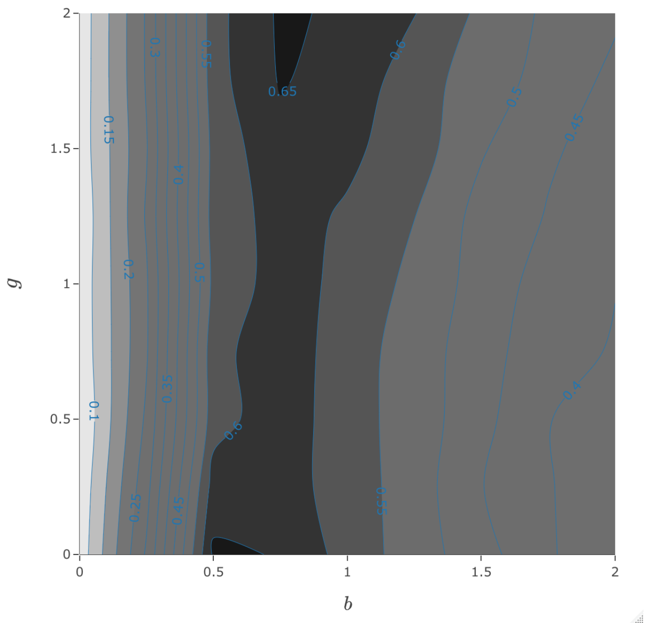
Figure 28: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(a) with $n = 200$, $p = 250$, $\rho = 0.25$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



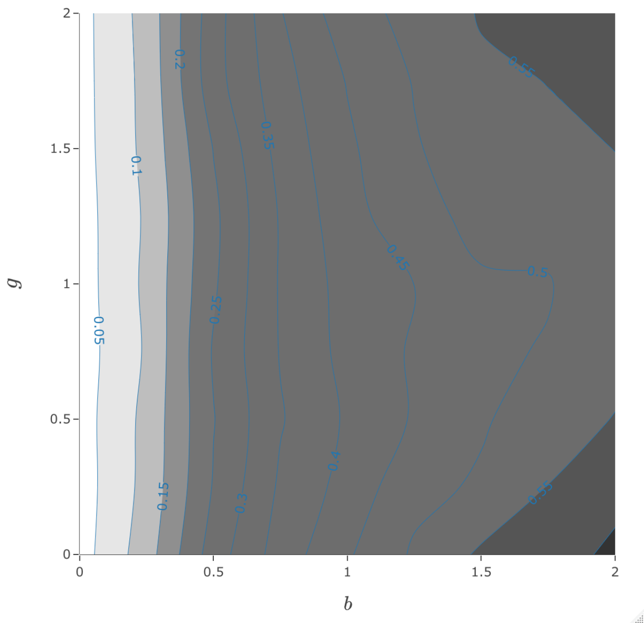
(a) Triple selection



(b) Poor man's approach

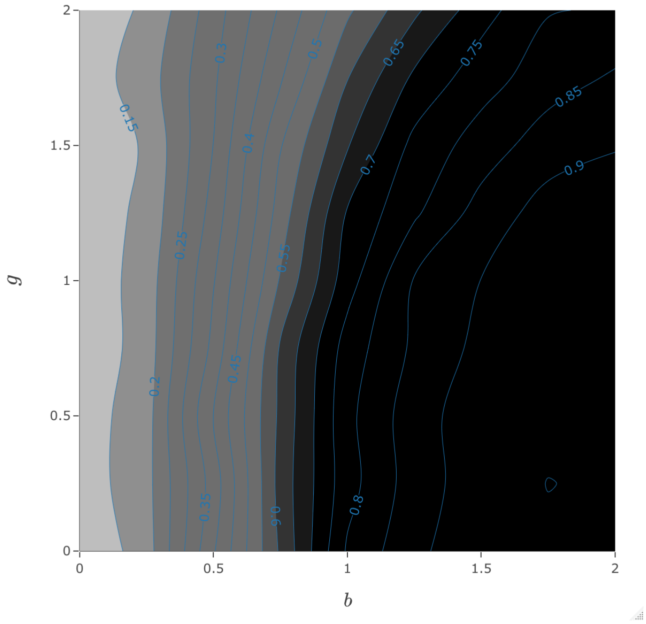


(c) Post-Lasso

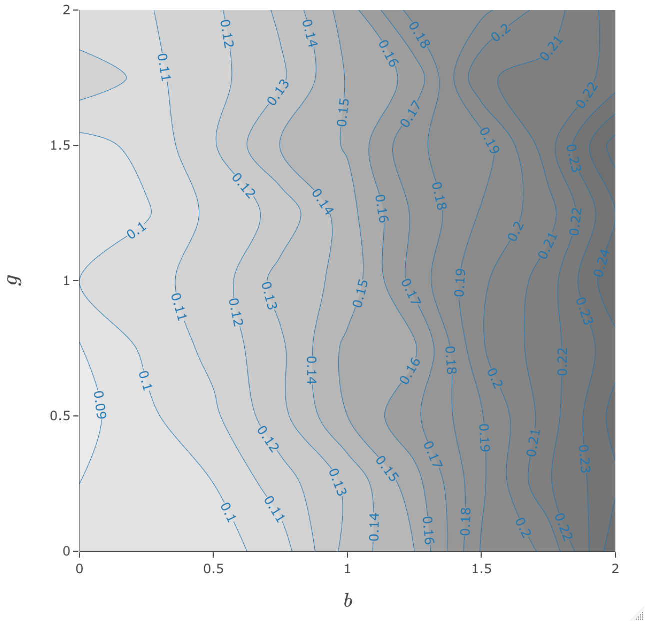


(d) Decorrelated test in Fang et al. (2017)

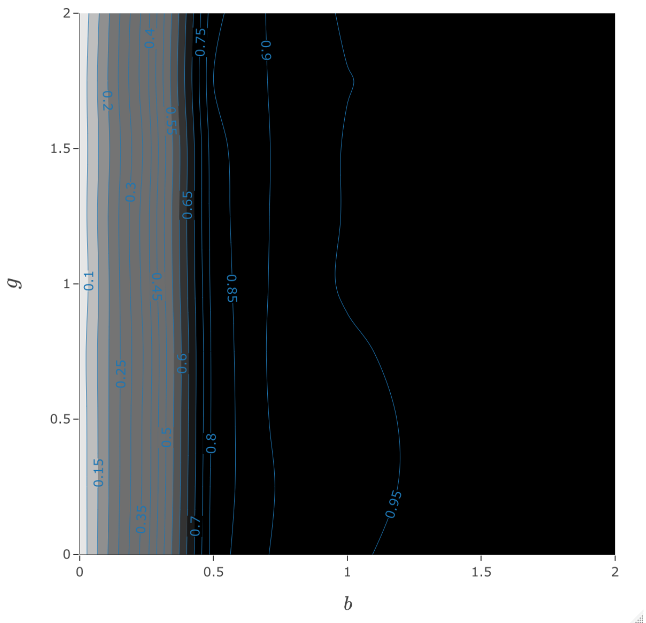
Figure 29: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(b) with $n = 200$, $p = 250$, $c_A = 1$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.



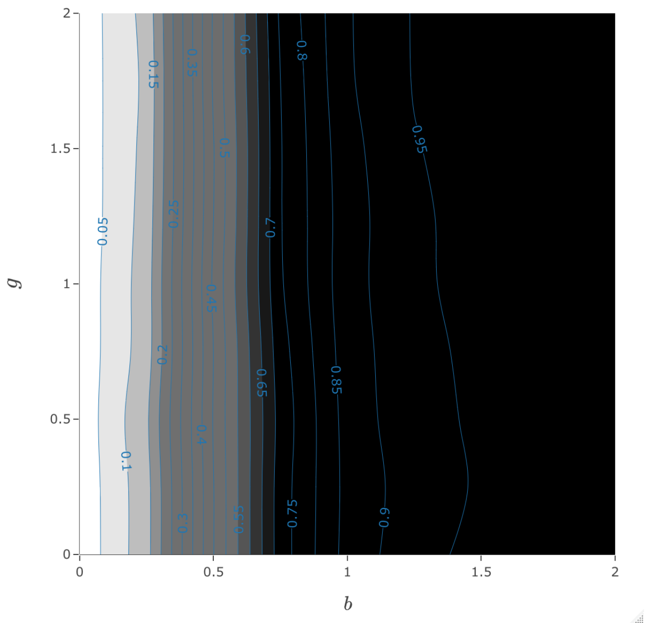
(a) Triple selection



(b) Poor man's approach



(c) Post-Lasso



(d) Decorrelated test in Fang et al. (2017)

Figure 30: Empirical Type I error rate at the 5% significance level of the different tests under Setting 2(b) with $n = 200$, $p = 250$, $c_A = 2$, $\eta_1 = 2$, $\beta_0 = 0$ and $\eta_0 = 0$.

References

Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica* **80,** 2369–2429.

Belloni, A. and Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli* **19,** 521–547.

Belloni, A., Chernozhukov, V., and Wei, Y. (2016). Post-selection inference for generalized linear models with many controls. *Journal of Business & Economic Statistics* **34,** 606–619.

Bertin, K., Le Pennec, E., and Rivoirard, V. (2011). Adaptive dantzig density estimation. *Ann. Inst. H. Poincaré Probab. Statist.* **47,** 43–74.

Bühlmann, P. and van de Geer, S. (2011). *Statistics for high-dimensional data*. Springer Series in Statistics. Springer, Heidelberg. Methods, theory and applications.

de la Peña, V., Lai, T., and Shao, Q.-M. (2009). *Self-normalized processes. Limit theory and statistical applications*. Probability and Its Applications. Springer Berlin Heidelberg.

Dukes, O., Avagyan, V., and Vansteelandt, S. (2020). Doubly robust tests of exposure effects under high-dimensional confounding. *Biometrics* .

Fang, E. X., Ning, Y., and Liu, H. (2017). Testing and confidence intervals for high dimensional proportional hazards models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **79,** 1415–1437.

Huang, J., Sun, T., Ying, Z., Yu, Y., and Zhang, C.-H. (2013). Oracle inequalities for the lasso in the cox model. *Ann. Statist.* **41,** 1142–1165.

Kasy, M. (2019). Uniformity and the delta method. *Journal of Econometric Methods* **8,** 20180001.

Lemler, S. (2016). Oracle inequalities for the lasso in the high-dimensional aalen multiplicative intensity model. *Ann. Inst. H. Poincaré Probab. Statist.* **52,** 981–1008.

Van der Vaart, A. W. (1998). *Asymptotic Statistics*. Asymptotic Statistics. Cambridge University Press.

von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the rth Absolute Moment of a Sum of Random Variables, $1r2$. *The Annals of Mathematical Statistics* **36,** 299 – 303.