arXiv:2302.00331v2 [stat.ME] 2 Feb 2023

双重加性治愈生存模型中的外生时变协变量及其在生育研究中的应用

Philippe Lambert*†和 Michaela Kreyenfeld‡

2023 年 1 月 31 日

摘要

扩展的治愈生存模型能够区分影响事件发生概率（或长期生存）的协变量与仅影响事件时间（或短期生存）的协变量。我们提出将有界累积风险模型推广至同时处理对长期和短期生存产生影响的时变（外生）协变量加性项。在这一框架中，惩罚参数的选择颇具挑战。我们提出一种基于贝叶斯 P 样条模型中拉普拉斯近似的快速算法。本方法的研究动机源于生育研究：在个体随访期间，女性的就业状态和收入等特征可能以复杂且频繁的方式发生变化。此外，我们利用德国养老保险基金的登记数据对该方法进行了说明，并据此研究女性时变收入与首次生育转变之间的关系。

关键词：加性模型；有界风险；治愈生存模型；生育研究；拉普拉斯近似；P样条；时变协变量。

1 引言

比例风险模型广泛用于分析事件发生时间数据及其与协变量的关联。它们能够通过假定风险比在时间上保持恒定来概括组间差异。治愈生存模型（Boag，1949年；Berkson 和 Gage，1952年）明确承认，所研究人群中有一定比例的人无论随访时间多长，都不会发生感兴趣的事件。检查估计的生存函数（如 Kaplan-Meier 曲线）可以揭示或证实这一点：当随访时间较长时，生存函数会趋于非零水平的平台。本文将重点讨论促发时间 （治愈）生存模型，该模型也称为有界累积风险模型（Yakovlev 和 Tsodikov，1996年；Tsodikov，1998年；Chen

*比利时列日大学数学研究所。电子邮箱：p.lambert@uliege.be
†比利时鲁汶天主教大学统计学、生物统计学与精算科学研究所（ISBA）。

‡德国柏林赫蒂学院。

等，1999年）。设$\mathbf{v} = (\mathbf{z},\mathbf{x})$，其中$\mathbf{z}$ 表示由分类协变量构成的 p–向量，$\mathbf{x}$表示由定量协变量构成的 J–向量（其中$\mathbf{0}$ 通常用作其参考值）。若 $S_p(t|\mathbf{v})$ 是具有这些特征的受试者（包括已治愈个体）的条件生存函数，则

$$S_p(t|\mathbf{v}) = \exp\{-\vartheta(\mathbf{v})F(t)\} \tag{1}$$

其中，$t > 0, \vartheta(\mathbf{v}) > 0$和 $F(t)$ 是一个累积分布函数，满足 $F(0) = 0$且 $F(T) = 1$，其中 T表示受试者被判定为治愈个体之前的最短生存时间。由$\mathbf{v}$ 定义的亚群中治愈个体的比例为

$$\pi(\mathbf{v}) = S_p(T|\mathbf{v}) = \exp\{-\vartheta(\mathbf{v})F(T))\} = \exp\{-\vartheta(\mathbf{v})\} > 0.$$

令 $f(t) = dF(t)/dt$。当 $\vartheta(\mathbf{v}) = \exp\{\eta_\vartheta(\mathbf{v})\}$且 $\eta_\vartheta(\mathbf{0}) = \beta_0$时，(1)对应于一个比例风险模型，其基线风险函数为 $e^{\beta 0} f(t)$，累积风险函数有界于 $\vartheta(\mathbf{v})$。风险函数的动态变化由 $f(t)$ 控制，不受协变量影响：它保证了比例风险模型的风险比恒定这一特征。这是一项关键假设，但并非总能得到恰当评估，从而可能影响研究结论的质量。即使比较组的估计生存曲线没有交叉，且在被右删失打断的随访期间，底层累积风险函数对数的平行性未受到质疑，比例风险假设仍可能在更长时间跨度上遭到违反。事实上，超过最大观测时间后，生存概率可能逐渐分化或趋于相似，甚至在更长时间后趋于一致。在后一种情况下，基于右删失数据报告的治疗组风险较低，仅表明事件发生延迟，而非治疗带来了长期获益。这促使人们提出扩展促发时间模型（Bremhorst and Lambert, 2016年；Bremhorst et al., 2016年）

$$S_p(t|\mathbf{v}, \tilde{\mathbf{v}}) = \exp\{-\vartheta(\mathbf{v})F(t|\tilde{\mathbf{v}})\}. \tag{2}$$

其中，总体风险函数的动态变化会随˜$\mathbf{v}$中的协变量而改变。还可以进一步考虑加速失效时间（AFT）模型或比例风险（PH）模型，以描述 F 对 ˜$\mathbf{v}$的依赖关系。

在这一框架中纳入时变协变量（TVC）颇具挑战。Dirick et al.（2019）在混合治愈模型框架下研究了这一问题，其中时变协变量仅限于非治愈个体的条件生存模型，因此，治愈概率的逻辑回归子模型中只能纳入恒定协变量。Lambert and Bremhorst（2020）在扩展促发时间模型框架下提出了更一般的建模形式，其中分类时变协变量不仅影响非治愈个体的事件发生时间，还纳入治愈概率的回归模型。然而，出于治愈概率子模型可识别性的考虑，每次协变量变化后的随访时长都必须足够长。

本文提出并研究了扩展促发时间模型的一种重新表述，使其能够处理分类或连续时变协变量中无限次的变化。此外，还可以在长期和短期生存子模型中联合考虑（恒定或时变）定量协变量非线性效应的加性项。

本文安排如下。第2节详细回顾扩展促发时间模型，并讨论如何利用 P样条纳入加性项。本文的方法学核心见第 3 节，其中提出了一种新方法，将分类或定量时变协变量纳入治愈生存模型。第 3.2 节介绍用于探索模型参数联合后验分布并计算其后验众数（最大后验估计，MAP）的算法。本文还提出了一种策略，用于选择调节长期和短期生存子模型中未知函数平滑度的惩罚参数，并给出一种易于实现且收敛迅速的算法。第4节通过一项大规模模拟研究评估该方法的优点。第 5 节通过分析养老金登记数据，以及研究德国女性时变收入与生育转变之间的关联，展示该方法的应用。最后，第6节对全文进行讨论。

2 扩展的促发时间模型

Bremhorst and Lambert（2016）首次探讨了在促发时间模型的非参数基线风险函数中使用协变量来改变其动态特征，见（2）：对 $\vartheta(\mathbf{v})$ 使用对数线性模型，对 $F(t|\tilde{\mathbf{v}})$ 使用比例风险模型。更具体地说，考虑后一个表达式的以下形式：$F(t|\tilde{\mathbf{v}}) = 1 - S_0(t)^{\exp(\eta_F(\tilde{\mathbf{v}}))}$ 其中，$S_0(t) = 1 - F(t|\mathbf{0})$ 是基线生存函数，$\eta_F(\cdot)$ 是协变量的一个（可能为非线性的）函数，并带有识别约束，例如 $\eta_F(\mathbf{0}) = 0$。于是，与（2）相关的总体累积风险函数和风险函数分别为：

$$H_p(t|\mathbf{v},\tilde{\mathbf{v}}) = \vartheta(\mathbf{v})F(t|\tilde{\mathbf{v}}) = \exp\{\eta_\vartheta(\mathbf{v})\}\left(1 - S_0(t)^{\exp(\eta_F(\tilde{\mathbf{v}}))}\right); \tag{3}$$

$$h_p(t|\mathbf{v},\tilde{\mathbf{v}}) = \vartheta(\mathbf{v})f(t|\tilde{\mathbf{v}}) = \mathrm{e}^{\eta_\vartheta(\mathbf{v})+\eta_F(\tilde{\mathbf{v}})} f_0(t)S_0(t)^{\exp(\eta_F(\tilde{\mathbf{v}}))-1}, \tag{4}$$

其中 $f_0(t) = -dS_0(t)/dt$。只要随访时间足够长，即使在某些协变量同时属于 v 和 ˜v 这一颇具挑战性的情形下，也可以解决可识别性问题；更多详情见 Lambert and Bremhorst（2019）。对于短期生存协变量的给定固定值 ˜v，（4）定义了一个比例风险模型，其治愈比例为 $\exp(-\vartheta(\mathbf{v}))$，且对于 v 的不同取值所对应的对比，风险比不随时间变化，即 $h_p(t|\mathrm{v2}, \tilde{\mathrm{v}})/h_p(t|\mathrm{v1}, \tilde{\mathrm{v}}) = \vartheta(\mathrm{v}_2)/\vartheta(\mathrm{v}_1)$。当针对 ˜v 的不同取值比较风险时，情况则不再如此，因为 $h_p(t|\mathrm{v}\ \tilde{\mathrm{v}}_2)/h_p(t|\mathrm{v}\ \tilde{\mathrm{v}}_1) = \tilde{\vartheta}(\tilde{\mathrm{v}}_2)/\tilde{\vartheta}(\tilde{\mathrm{v}}_1)S_0(t)^{\vartheta(\tilde{\mathrm{v}}_2)-\vartheta(\tilde{\mathrm{v}}_1)}$（其中 $\tilde{\vartheta}(\tilde{\mathrm{v}}) = \mathrm{e}^{\eta_F}(\tilde{\mathrm{v}})$）会随时间变化。

Bremhorst and Lambert（2016年）考虑使用 B 样条的线性组合来指定 $h_0(t) = f_0(t)/S_0(t)$。这里，优先采用基于 P 样条（Eilers and Marx，1996年）的灵活形式来表示 $f_0(\cdot)$，

$$f_0(t) = \frac{\exp\left(\sum_k b_k(t)\phi_k\right)}{\int_0^T \exp\left(\sum_k b_k(u)\phi_k\right) du}\mathbb{1}_{\{0\le t\le T\}} \tag{5}$$

其中，$\{b_k(\cdot): \ k = 1, \ldots, K\}$ 表示与区间 $(0, T)$ 上等距结点相对应的大型 B 样条基，而 $\boldsymbol{\phi} = (\phi_k)_{k=1}^K$ 是由样条参数构成的向量，满足 $\phi_{\lfloor K/2 \rfloor} = 0$（用于可识别性）。它与参考人群风险函数直接相关：$h_p(t|\mathbf{0}, \mathbf{0}) = \mathrm{e}^{\beta_0} f_0(t)$，其中 β_0 决定总风险暴露，$f_0(t)$则描述其随时间的分布。通过惩罚样条系数的变化来保证 $f_0(\cdot)$的平滑度，见第 3.2 节。连续协变量（如生育研究中年龄或收入对妊娠概率及其发生时间的影响）可能产生非线性效应，这些效应将采用加性形式建模。假设观测到 n 个独立个体，其数据为$\mathcal{D} = \cup_{i=1}^n \mathcal{D}_i$，其中 $\mathcal{D}_i = \{t_i, \delta_i, \mathbf{v}_i, \tilde{\mathbf{v}}_i\}$, δ_i 是随访时间t_i 的事件指示变量，而 $\mathbf{v}_i$, $\tilde{\mathbf{v}}_i$ 是个体 i的长期和短期生存协变量。我们通过采用非线性形式描述定量协变量对 $\eta_\vartheta(\mathbf{v})$和 $\eta_F(\tilde{\mathbf{v}})$的影响，为扩展促发时间模型增加灵活性，

$$\left(\eta_\vartheta(\mathbf{v}_i)\right)_{i=1}^n = \left(\beta_0 + \sum_{k=1}^p \beta_k z_{ik} + \sum_{j=1}^J f_j(x_{ij})\right)_{i=1}^n = \mathbf{Z}\boldsymbol{\beta} + \sum_{j=1}^J \mathbf{f}_j \ , \tag{6}$$

$$\left(\eta_F(\tilde{\mathbf{v}}_i)\right)_{i=1}^n = \left(\sum_{k=1}^{\tilde{p}} \gamma_k \tilde{z}_{ik} + \sum_{j=1}^{\tilde{J}} \tilde{f}_j(\tilde{x}_{ij})\right)_{i=1}^n = \tilde{\mathbf{Z}}\boldsymbol{\gamma} + \sum_{j=1}^{\tilde{J}} \tilde{\mathbf{f}}_j \ , \tag{7}$$

其中，$f_j(\cdot)$ 和 $\tilde{f}_j(\cdot)$ 分别表示平滑加性项，用于量化相关定量协变量对长期生存和短期生存的影响；$\mathbf{f}_j = \left(f_j(x_{ij})\right)_{i=1}^n$ 和 $\tilde{\mathbf{f}}_j = \left(\tilde{f}_j(\tilde{x}_{ij})\right)_{i=1}^n$ 表示将个体对应的取值堆叠成向量后的结果；$\mathbf{Z}$ 是$n \times (1 + p)$ 设计矩阵，其中包含一列全为 1 的截距项，以及每个其他分类协变量对应的一列；$n \times \tilde{p}$ 设计矩阵 $\tilde{\mathbf{Z}}$ 的构造方式类似（但由于没有截距项，不包含全为 1 的列）。现在，考虑由 $(L + 1)$ 个三次B样条 $\{s_{j\ell}^*(\cdot)\}_{\ell=1}^{L+1}$ 构成的基，这些样条对应于在x_j取值范围 $(x_j^{\min}, x_j^{\max})$ 上等间距分布的节点。为确保加性模型的可识别性，使用$s_{j\ell}(\cdot) = s_{j\ell}^*(\cdot) - \frac{1}{x_j^{\max} - x_j^{\min}} \int_{x_j^{\min}}^{x_j^{\max}} s_{j\ell}^*(u)du \, (\ell = 1, \ldots, L)$ 对其重新中心化。对与短期生存相关的协变量也采用类似处理，得到重新中心化的B样条，记为 $\tilde{s}_{j\ell}(\cdot)$。随后，条件长期生存和短期生存子模型中的加性项可近似表示为此类（重新中心化）B样条的线性组合：$\mathbf{f}_j = \left(\sum_{\ell=1}^L s_{j\ell}(x_{ij})\theta_{\ell j}\right)_{i=1}^n = \mathbf{S}_j\boldsymbol{\theta}_j$, $\tilde{\mathbf{f}}_j = \left(\sum_{\ell=1}^L \tilde{s}_{j\ell}(\tilde{x}_{ij})\tilde{\theta}_{\ell j}\right)_{i=1}^n = \tilde{\mathbf{S}}_j\tilde{\boldsymbol{\theta}}_j$，其中$[\mathbf{S}_j]_{i\ell} = s_{j\ell}(x_{ij})$、$[\tilde{\mathbf{S}}_j]_{i\ell} = \tilde{s}_{j\ell}(\tilde{x}_{ij})$、$\left(\boldsymbol{\theta}_j\right)_\ell = \theta_{\ell j}$ 和 $\left(\tilde{\boldsymbol{\theta}}_j\right)_\ell = \tilde{\theta}_{\ell j}$。因此，采用向量记法，式（6）和（7）中的条件长期生存和短期生存线性预测因子可改写为 $\left(\eta_{\vartheta i} = \eta_\vartheta(\mathbf{v}_i)\right)_{i=1}^n = \mathbf{X}\boldsymbol{\psi}$, $\left(\eta_{Fi} = \eta_F(\tilde{\mathbf{v}}_i)\right)_{i=1}^n = \tilde{\mathbf{X}}\tilde{\boldsymbol{\psi}}$，其中设计矩阵为 $\mathbf{X} = [\mathbf{Z}, \mathbf{S}1, \ldots, \mathbf{S}_J] = [\mathbf{Z}, \mathbf{S}] \in \mathbb{R}^{n \times q}$, $\tilde{\mathbf{X}} = [\tilde{\mathbf{Z}}, \tilde{\mathbf{S}}1, \ldots, \tilde{\mathbf{S}}_J] = [\tilde{\mathbf{Z}}, \tilde{\mathbf{S}}] \in \mathbb{R}^{n \times \tilde{q}}$；样条参数矩阵（每个加性项对应一列）为$\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_J]$ in $\mathbb{R}^{L \times J}$, $\tilde{\boldsymbol{\Theta}} = [\tilde{\boldsymbol{\theta}}_1, \ldots, \tilde{\boldsymbol{\theta}}_{\tilde{J}}]$ in $\mathbb{R}^{L \times \tilde{J}}$；（堆叠的）回归参数向量为 $\boldsymbol{\psi} = \left(\boldsymbol{\beta}, \mathrm{vec}(\boldsymbol{\Theta})\right)$ in $\mathbb{R}^q$, $\tilde{\boldsymbol{\psi}} = \left(\boldsymbol{\gamma}, \mathrm{vec}(\tilde{\boldsymbol{\Theta}})\right)$ in $\mathbb{R}^{\tilde{q}}$，其中 $q = (1 + p + JL)$ 和 $\tilde{q} = (\tilde{p} + \tilde{J}L)$。注意，式（6）对应于治愈概率的加性对数-对数模型，或者等价地，长期事件的互补对数-对数模型

概率。

3 带时变协变量的治愈模型

3.1 模型设定

现在假设协变量是外生的，且其取值会随时间变化。我们提出在总体层面使用以下方式对风险率建模：

$$\begin{aligned} h_p(t|\mathbf{v}(t),\tilde{\mathbf{v}}(t)) &= \vartheta(\mathbf{v}(t))f(t|\tilde{\mathbf{v}}(t)) \\ &= \mathrm{e}^{\eta_\vartheta(\mathbf{v}(t))+\eta_F(\tilde{\mathbf{v}}(t))} f_0(t) S_0(t)^{\exp(\eta_F(\tilde{\mathbf{v}}(t))-1}, \end{aligned} \tag{8}$$

由此得到一种带时变协变量的（促发时间）治愈生存模型，简称 TVcure 模型。相关的累积风险函数可通过积分进行数值计算。当协变量为常数这一特殊情形下，我们可得到 Bremhorst and Lambert （ 2016 ）中人口风险和累积风险函数的表达式，参见第 2 节中的（3）和（4），其参数解释已在前文讨论。一般情况下，线性预测因子 η_ϑ 和 η_F 不仅随个体变化，也可能随时间变化。因此，相关设计矩阵也取决于时间，即 $\left(\eta_\vartheta(\mathbf{v}_i(t))\right)_{i=1}^n = \boldsymbol{\mathcal{X}}_t\boldsymbol{\psi}$, $\left(\eta_F(\tilde{\mathbf{v}}_i(t))\right)_{i=1}^n = \tilde{\boldsymbol{\mathcal{X}}}_t\tilde{\boldsymbol{\psi}}$。

进一步假设，每个个体的数据都可以按固定时间间隔（以 dt 个时间单位计）定期报告，例如，从年龄20岁（$t = 0$）起，直到事件发生时（如首次妊娠）或随访期结束时（如果到那时她仍未生育），按月报告（$dt = 1$ 个月）女性状态及其协变量值。于是，第 i个个体的数据形式如下： $\mathcal{D}_i = \{(d_{it}, \mathbf{v}_i(t), \tilde{\mathbf{v}}_i(t)) : t = 1, \ldots\ldots, t_i\}$，其中 d_{it} 是事件指示变量，对于所有 t均恒等于0；但最后一个值 d_{it_i} 可能等于1，表示在（ $t_i - dt$， t_i）内观察到事件（ $\delta_i = 1$ ），否则为0。

3.2 推断

假设协变量在一个时间单位 dt内保持不变，则在时间 $t - dt$ 仍处于风险中的第 i个受试者的 d_{it} 条件分布仍近似服从均值为 $\mu_{it} = h_p(t|\mathbf{v}(t), \tilde{\mathbf{v}}(t))\, dt$的泊松分布（Lindsey, 1995年）。该受试者的对数似然贡献为

$$\ell_i(\boldsymbol{\phi}, \boldsymbol{\psi}, \tilde{\boldsymbol{\psi}}|\mathcal{D}_i) = -\sum_{t=1}^{t_i} \mu_{it} + d_{it_i} \log \mu_{it_i},$$

并在式（5）至（8）中明确了 μ_{ij} 对三个参数向量的依赖关系。为样条参数设置平滑先验，以弥补模型描述中采用大规模 B 样条基所带来的灵活性；这些基用于指定

$f_0(t)$ 以及式（6）和（7）中的加性项（Fahrmeir 和 Lang，2001年），

$$
\begin{aligned}
p(\boldsymbol{\phi}|\tau_0) &\propto \exp\left(-\frac{1}{2}\,\boldsymbol{\phi}^{\mathsf{T}}(\tau_0\mathbf{P}_0)\,\boldsymbol{\phi}\right)\ , \\
p(\boldsymbol{\theta}_j|\lambda_j) &\propto \exp\left(-\frac{1}{2}\,\boldsymbol{\theta}_j^{\mathsf{T}}(\lambda_j\mathbf{P})\,\boldsymbol{\theta}_j\right)\quad (j=1,\ldots,J)\ , \\
p(\tilde{\boldsymbol{\theta}}_j|\tilde{\lambda}_j) &\propto \exp\left(-\frac{1}{2}\,\tilde{\boldsymbol{\theta}}_j^{\mathsf{T}}(\tilde{\lambda}_j\tilde{\mathbf{P}})\,\tilde{\boldsymbol{\theta}}_j\right)\quad (j=1,\ldots,\tilde{J}).
\end{aligned}
\tag{9}
$$

对惩罚参数采用伽马先验，$\tau_0,\lambda_j,\tilde{\lambda}_j \sim \mathcal{G}\left(1,d=10^{-4}\right)$（Lang and Brezger，2004年）；或采用伽马混合分布（Jullion and Lambert，2007年），其特殊情形包括对这些参数的平方根采用半柯西先验（Lambert and Bremhorst，2019年）。假设与其他协变量 $\mathbf{z}$ 和 $\tilde{}\mathbf{z}$相关的参数服从联合正态先验，$\boldsymbol{\beta}\sim\mathcal{N}\left(\breve{b},\mathbf{Q}^{-1}\right)$，$\boldsymbol{\gamma}\sim\mathcal{N}\left(\breve{}\mathbf{g}\ \ \tilde{}\mathbf{Q}^{-1}\right)$，则回归参数和$\tilde{}$样条参数在 $\boldsymbol{\psi}$ 和 $\boldsymbol{\psi}$ 中的联合先验会诱导出高斯马尔可夫随机场（GMRF）（Rue and Held，2005年），因为它们可以写成

$$
\begin{aligned}
p(\boldsymbol{\psi}|\boldsymbol{\lambda}) &\propto \exp\left(-\frac{1}{2}\ (\boldsymbol{\psi}-\mathbf{b})^{\mathsf{T}}\mathbf{K}_\lambda(\boldsymbol{\psi}-\mathbf{b})\right)\ ; \\
p(\tilde{\boldsymbol{\psi}}|\tilde{\boldsymbol{\lambda}}) &\propto \exp\left(-\frac{1}{2}\ (\tilde{\boldsymbol{\psi}}-\mathbf{g})^{\mathsf{T}}\tilde{\mathbf{K}}_{\tilde{\lambda}}(\tilde{\boldsymbol{\psi}}-\mathbf{g})\right),
\end{aligned}
\tag{10}
$$

其中 $\mathbf{b}=(\breve{}\mathbf{b}^{\mathsf{T}},\mathbf{0}_{JL}^{\mathsf{T}})^{\mathsf{T}}$、$\mathrm{K}_\lambda=\mathrm{diag}(\mathrm{Q}$、$\mathrm{P}_\lambda)$、$\mathrm{P}_\lambda=\boldsymbol{\Lambda}\otimes\mathrm{P}$、$[\boldsymbol{\Lambda}]_{jj'}=\delta_{jj'}\lambda_j$、$\mathbf{g}=(\breve{}\mathbf{g}^{\mathsf{T}},\mathbf{0}_{\tilde{J}L}^{\mathsf{T}})^{\mathsf{T}}$、$\tilde{}\mathrm{K}\tilde{\lambda}=\mathrm{diag}(\tilde{}\mathrm{Q}\ \ \tilde{}\mathrm{P}_{\tilde{\lambda}})$、$\tilde{}\mathrm{P}_\lambda=\tilde{\boldsymbol{\Lambda}}\otimes\tilde{}\mathrm{P}$ 以及 $[\tilde{\boldsymbol{\Lambda}}]_{jj'}=\delta_{jj'}\tilde{\lambda}_j$。设 $\boldsymbol{\lambda}=(\lambda_j)_{j=1}^{J}$、$\tilde{\boldsymbol{\lambda}}=(\tilde{\lambda}_j)_{j=1}^{\tilde{J}}$，其联合先验分布为$\tilde{}$ $\tilde{}(p)\,\boldsymbol{\lambda}(\propto\prod_j p)$、$\lambda_j(p)\,\boldsymbol{\lambda}(\propto\prod_j p)$。若 λ_j 表示对数似然，则模型参数的联合后验分布可直接由贝叶斯定理得出，

$$
\begin{aligned}
&p(\boldsymbol{\phi},\boldsymbol{\psi},\tilde{\boldsymbol{\psi}},\tau_0,\boldsymbol{\lambda},\tilde{\boldsymbol{\lambda}}|\mathcal{D})\propto \\
&\qquad \exp\{\ell(\boldsymbol{\phi},\boldsymbol{\psi},\tilde{\boldsymbol{\psi}}|\mathcal{D})\}\,p(\boldsymbol{\phi}|\tau_0)p(\tau_0)\,p(\boldsymbol{\psi}|\boldsymbol{\lambda})p(\boldsymbol{\lambda})\,p(\tilde{\boldsymbol{\psi}}|\tilde{\boldsymbol{\lambda}})p(\tilde{\boldsymbol{\lambda}}).
\end{aligned}
$$

3.2.1 回归参数的条件估计

回归参数和样条参数的条件后验众数，$\boldsymbol{\zeta}=(\boldsymbol{\psi}^{\mathsf{T}},\tilde{\boldsymbol{\psi}}^{\mathsf{T}})^{\mathsf{T}}$和 $\boldsymbol{\phi}$，与其条件惩罚最大似然估计（PMLE）一致，优化目标为

$$
\ell_p=\ell-\frac{\tau_0}{2}\,\boldsymbol{\phi}^{\mathsf{T}}\mathbf{P}_0\,\boldsymbol{\phi}-\frac{1}{2}\ (\boldsymbol{\psi}-\mathbf{b})^{\mathsf{T}}\mathbf{K}_\lambda(\boldsymbol{\psi}-\mathbf{b})-\frac{1}{2}\ (\tilde{\boldsymbol{\psi}}-\mathbf{g})^{\mathsf{T}}\tilde{\mathbf{K}}_{\tilde{\lambda}}(\tilde{\boldsymbol{\psi}}-\mathbf{g})\,.
$$

可通过牛顿-拉夫森（N-R）算法迭代求得；该算法基于各自梯度和精度矩阵的显式形式构建，详见附录 A。实际操作中，给定 $\boldsymbol{\phi}$ 和惩罚参数 $\boldsymbol{\tau}=(\boldsymbol{\lambda}^{\mathsf{T}},\ \tilde{\boldsymbol{\lambda}}^{\mathsf{T}})^{\mathsf{T}}$的取值，重复以下替代步骤直至收敛：

$$
\boldsymbol{\zeta}\longleftarrow\boldsymbol{\zeta}-\mathcal{H}_\tau^{-1}\mathbf{U}_\tau^{\zeta}\ .
\tag{11}
$$

以类似方式迭代求得式 (5) 中定义 $f_0(t)$ 的样条参数 $\boldsymbol{\phi}$ 的估计值，迭代时以惩罚参数 τ_0为条件，

$$\boldsymbol{\phi}_{-k} \longleftarrow \boldsymbol{\phi}_{-k} - \left((\boldsymbol{H}^{\phi\phi}_{\tau_0})_{-k,-k}\right)^{-1} (\mathbf{U}^{\phi}_{\tau_0})_{-k} \ , \tag{12}$$

为确保可识别性，将其中一个向量分量任意设为零，$\phi_k = 0$。

3.2.2 惩罚参数的选择

用于调节加性项平滑度的惩罚参数 $\boldsymbol{\tau} = (\boldsymbol{\lambda}^{\intercal},\ \tilde{\boldsymbol{\lambda}}^{\intercal})^{\intercal}$ 的边际后验分布可利用以下恒等式得到（隐含地以 $\boldsymbol{\phi}$ 和 τ_0为条件），

$$p(\boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}|\mathcal{D}) = p(\boldsymbol{\psi}, \tilde{\boldsymbol{\psi}}, \boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}|\mathcal{D}) / p(\boldsymbol{\psi}, \tilde{\boldsymbol{\psi}}|\boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}, \mathcal{D}) \ ,$$

在分母中以拉普拉斯近似替代样条参数的条件后验分布；非参数双加性位置尺度模型中类似的策略可参见 Lambert（2021年）。在条件后验众数处计算该表达式，可得到惩罚参数边际后验分布的以下近似，

$$p(\boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}|\mathcal{D}) \propto p(\hat{\boldsymbol{\psi}}_{\lambda}, \hat{\tilde{\boldsymbol{\psi}}}_{\tilde{\lambda}}, \boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}|\mathcal{D}) \left|\Sigma_{\tau}^{-1}\right|^{-1/2} \ , \tag{13}$$

其中，精度矩阵中的分块

$$\Sigma_{\tau}^{-1} = -\mathcal{H}_{\tau} = \begin{bmatrix} -\mathbf{H}^{\psi\psi}_{\lambda} & -\mathbf{H}^{\psi\tilde{\psi}} \\ -\mathbf{H}^{\tilde{\psi}\psi} & -\mathbf{H}^{\tilde{\psi}\tilde{\psi}}_{\tilde{\lambda}} \end{bmatrix} \ , \tag{14}$$

具有显式形式，见附录 A。最大化该边际后验分布即可选择惩罚参数。注意，（10）中 $\boldsymbol{\psi}$ 和 $\boldsymbol{\psi}$ 的~先验分布的归一化常数取决于惩罚参数，因此（13）的对数可写为（略去一个加性常数）

$$\begin{aligned} \log p(\boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}|\mathcal{D}) &\doteq \ell_p(\hat{\boldsymbol{\phi}}_{\tau_0}, \hat{\boldsymbol{\zeta}}_{\tau}|\tau, \mathcal{D}) + \frac{1}{2}\left\{\log|\mathbf{K}_{\lambda}|^{+} + \log|\tilde{\mathbf{K}}_{\tilde{\lambda}}|^{+} - \log|-\mathcal{H}_{\tau}|\right\} \\ &= \ell_p(\hat{\boldsymbol{\phi}}_{\tau_0}, \hat{\boldsymbol{\zeta}}_{\tau}|\tau, \mathcal{D}) + \frac{\rho(\mathbf{P})}{2}\sum_{j=1}^{J}\log\lambda_j + \frac{\rho(\tilde{\mathbf{P}})}{2}\sum_{j=1}^{\tilde{J}}\log\tilde{\lambda}_j - \frac{1}{2}\log|-\mathcal{H}_{\tau}| \ , \end{aligned} \tag{15}$$

其中，$|\mathbf{A}|^{+}$ 表示半正定矩阵 $\mathbf{A}$ 的非零特征值之积。设 $\Sigma_{\tau}(\boldsymbol{\theta}_j)$ 为 Σ_{τ} 中与 $(\boldsymbol{\psi}^{\intercal}, \tilde{\boldsymbol{\psi}}^{\intercal})^{\intercal}$ 里的子向量 $\boldsymbol{\theta}_j$ 相对应的 $L \times L$ 子矩阵。已知

$$\frac{\partial \ell_p(\hat{\boldsymbol{\phi}}_{\tau_0}, \hat{\boldsymbol{\zeta}}_{\tau}|\tau, \mathcal{D})}{\partial \lambda_j} = \frac{\partial \ell_p(\hat{\boldsymbol{\phi}}_{\tau_0}, \hat{\boldsymbol{\zeta}}_{\tau}|\tau, \mathcal{D})}{\partial \hat{\boldsymbol{\psi}}_{\lambda}^{\intercal}} \frac{\partial \hat{\boldsymbol{\psi}}_{\lambda}}{\partial \lambda_j} - \frac{1}{2}\hat{\boldsymbol{\theta}}_{j\lambda}^{\intercal}\mathbf{P}\hat{\boldsymbol{\theta}}_{j\lambda} \quad (1 \le j \le J)$$

该表达式中的第一个因子等于零，并且记住，对于正定矩阵 A_s， $\partial \log |A_s|/\partial s = \mathrm{Tr}\left(A_s^{-1}\partial A_s/\partial s\right)$，则有

$$\frac{\partial \log p(\boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}|\mathcal{D})}{\partial \lambda_j} = \frac{1}{2}\left\{\frac{\rho(\mathbf{P})}{\lambda_j} - \hat{\boldsymbol{\theta}}_{j\lambda}^{\mathsf{T}}\mathbf{P}\hat{\boldsymbol{\theta}}_{j\lambda} - \mathrm{Tr}\left(\Sigma_\tau(\boldsymbol{\theta}_j)\mathbf{P}\right)\right\}. \tag{16}$$

˜可以得到一个类似的表达式（其中包含 $1 \le j \le J$），用于

$$\frac{\partial \log p(\boldsymbol{\lambda}, \tilde{\boldsymbol{\lambda}}|\mathcal{D})}{\partial \tilde{\lambda}_j} = \frac{1}{2}\left\{\frac{\rho(\tilde{\mathbf{P}})}{\tilde{\lambda}_j} - \hat{\tilde{\boldsymbol{\theta}}}_{j\tilde{\lambda}}^{\mathsf{T}}\tilde{\mathbf{P}}\hat{\tilde{\boldsymbol{\theta}}}_{j\tilde{\lambda}} - \mathrm{Tr}\left(\Sigma_\tau(\tilde{\boldsymbol{\theta}}_j)\tilde{\mathbf{P}}\right)\right\}. \tag{17}$$

˜ $\boldsymbol{\lambda}$ 和 $\boldsymbol{\lambda}$ 的最大后验估计是将式 (16) 和 (17) 对所有 j 置为零后得到的解。可通过不动点法求解，迭代进行以下替换直至收敛：

$$\begin{aligned}
\lambda_j^{-1} &\longleftarrow \frac{\hat{\boldsymbol{\theta}}_{j\lambda}^{\mathsf{T}}\mathbf{P}\hat{\boldsymbol{\theta}}_{j\lambda} + \mathrm{Tr}\left(\Sigma_\tau(\boldsymbol{\theta}_j)\mathbf{P}\right)}{\rho(\mathbf{P})} = \frac{\mathbb{E}(\boldsymbol{\theta}_j^{\mathsf{T}}\mathbf{P}\boldsymbol{\theta}_j|\boldsymbol{\lambda}, \mathcal{D})}{\rho(\mathbf{P})} \qquad (1 \le j \le J) \\
\tilde{\lambda}_j^{-1} &\longleftarrow \frac{\hat{\tilde{\boldsymbol{\theta}}}_{j\tilde{\lambda}}^{\mathsf{T}}\tilde{\mathbf{P}}\hat{\tilde{\boldsymbol{\theta}}}_{j\tilde{\lambda}} + \mathrm{Tr}\left(\Sigma_\tau(\tilde{\boldsymbol{\theta}}_j)\tilde{\mathbf{P}}\right)}{\rho(\tilde{\mathbf{P}})} = \frac{\mathbb{E}(\tilde{\boldsymbol{\theta}}_j^{\mathsf{T}}\tilde{\mathbf{P}}\tilde{\boldsymbol{\theta}}_j|\tilde{\boldsymbol{\lambda}}, \mathcal{D})}{\rho(\tilde{\mathbf{P}})} \qquad (1 \le j \le \tilde{J}).
\end{aligned} \tag{18}$$

式 (18) 与条件期望之间的联系源于此前对 $(\boldsymbol{\theta}|\boldsymbol{\lambda}, \mathcal{D})$ 和 $(\boldsymbol{\theta}|\boldsymbol{\lambda}, \mathcal{D})$ 进行的拉普拉斯近似。结合式 (11) 和 (18)，可得到算法 1，用于选择 $\boldsymbol{\tau}$ 并估计 $\boldsymbol{\zeta}$ （给定 $\boldsymbol{\phi}$ 的取值）。

通过调节参数 τ_0 来调整 $f_0(t)$的平滑度，其选择过程类似。给定回归参数后，最大化 $p(\tau_0|\mathcal{D})$（采用与 $p(\boldsymbol{\lambda}, \boldsymbol{\lambda}|\mathcal{D})$ 相同类型的论证进行近似）可通过以下替换实现，并重复迭代直至收敛：

$$\tau_0^{-1} \longleftarrow \frac{\hat{\boldsymbol{\phi}}_{\tau_0}^{\mathsf{T}}\mathbf{P}_0\hat{\boldsymbol{\phi}}_{\tau_0} + \mathrm{Tr}\left((-\mathbf{H}_{\tau_0}^{\phi\phi})^{-1}\mathbf{P}_0\right)}{\rho(\mathbf{P}_0)} = \frac{\mathbb{E}(\boldsymbol{\phi}^{\mathsf{T}}\mathbf{P}_0\boldsymbol{\phi}|\tau_0, \mathcal{D})}{\rho(\mathbf{P}_0)} \tag{19}$$

结合（12）和（19），可得到算法2，用于选择 τ_0 并估计 $\boldsymbol{\phi}$ （给定 $\boldsymbol{\zeta}$的取值）。

3.2.3 全局估计算法

含时变协变量的双加性TVcure模型可使用算法 3 进行拟合。该算法交替进行以下步骤，直至收敛：选择并估计回归子模型中的参数$\{\boldsymbol{\tau}, \boldsymbol{\zeta}\}$，以及估计用于描述基线短期生存动态的参数 $\{\tau_0, \boldsymbol{\phi}\}$（分类协变量取参考值，加性项设为零）。算法可采用以下初始值：$\boldsymbol{\zeta}$ 和 $\boldsymbol{\phi}$中的所有样条参数和回归参数均设为0； $\boldsymbol{\tau}$ 和 τ_0中的所有惩罚参数均设为适中的较大值（例如100）。作者使用纯R代码实现了该程序，并将其纳入名为tvcure的软件包中。算法收敛迅速，在普通笔记本电脑上只需几秒钟。

算法 1： $\boldsymbol{\tau}$ 的选择和 $\boldsymbol{\zeta}$ 的估计（给定 $\boldsymbol{\phi}$）

输入： 短期生存子模型的样条参数 $\boldsymbol{\phi}$ 、数据 $\mathcal{D} = \cup_{i=1}^{n} \{(d_{it},\ \mathbf{v}_i(t)\ \tilde{\mathbf{v}}_i(t))\colon\ t = 1,\ \ldots\ldots,\ t_i\}$。

输出： 已选择的 $\boldsymbol{\tau} = (\boldsymbol{\lambda}^{\mathsf{T}},\ \tilde{\boldsymbol{\lambda}}^{\mathsf{T}})^{\mathsf{T}}$ 及估计的 $\hat{\boldsymbol{\zeta}}_{\tau}$（给定 $\boldsymbol{\phi}$）。

注： $\boldsymbol{\zeta} = (\boldsymbol{\psi}^{\mathsf{T}},\ \tilde{\boldsymbol{\psi}}^{\mathsf{T}})^{\mathsf{T}}$ 其中 $\boldsymbol{\psi} = (\boldsymbol{\beta},\ \mathrm{vec}(\boldsymbol{\Theta}))$, $\tilde{\boldsymbol{\psi}} = \left(\tilde{\boldsymbol{\gamma}}^{\mathsf{T}}, \mathrm{vec}(\tilde{\boldsymbol{\Theta}})^{\mathsf{T}}\right)$

重复

  重复

    使用 (21)、(22) 和 (23) 计算 $\mathbf{U}_{\tau}^{\zeta}$ 和 $\mathcal{H}_{\tau}$。

    更新 $\boldsymbol{\zeta} \longleftarrow \boldsymbol{\zeta} - \mathcal{H}_{\tau}^{-1}\mathbf{U}_{\tau}^{\zeta}$

  直到 $||\mathbf{U}_{\tau}^{\zeta}||_{\infty} < \epsilon$;

  重复

    $\lambda_j^{-1} \longleftarrow \left\{\boldsymbol{\theta}_j^{\mathsf{T}}\mathbf{P}\boldsymbol{\theta}_j + \mathrm{Tr}\left(\Sigma_{\tau}(\boldsymbol{\theta}_j)\mathbf{P}\right)\right\}/\rho(\mathbf{P}) \quad (1 \le j \le J)$

    $\tilde{\lambda}_j^{-1} \longleftarrow \left\{\tilde{\boldsymbol{\theta}}_j^{\mathsf{T}}\tilde{\mathbf{P}}\tilde{\boldsymbol{\theta}}_j + \mathrm{Tr}\left(\Sigma_{\tau}(\tilde{\boldsymbol{\theta}}_j)\ \tilde{\mathbf{P}}\right)\right\}/\rho(\tilde{\mathbf{P}})\ (1 \le j \le \tilde{J})$

    使用 (21)、(22) 和 (23) 更新 $\Sigma_{\tau} = (-\mathcal{H}_{\tau})^{-1}$。

  直到收敛;

直到收敛;

算法2： 选择 τ_0 并估计 $\boldsymbol{\phi}$（给定 $\boldsymbol{\zeta}$）

输入： 回归和样条参数 $\boldsymbol{\zeta}$, 数据 $\mathcal{D} = \cup_{i=1}^{n} \{(d_{it},\ \mathbf{v}_i(t)\ \tilde{\mathbf{v}}_i(t))\colon\ t = 1,\ \ldots\ldots,\ t_i\}$。

输出： 已选定 τ_0 并估计 $\hat{\boldsymbol{\phi}}_{\tau_0}$（给定 $\boldsymbol{\zeta}$）。

重复

  重复

    使用 (20) 计算 $\mathbf{U}_{\tau_0}^{\phi}$ 和 $\mathbf{H}_{\tau_0}^{\phi\phi}$

    用 $\phi_k = 0$ 更新 $\boldsymbol{\phi}_{-k} \longleftarrow \boldsymbol{\phi}_{-k} - \left((\boldsymbol{H}_{\tau_0}^{\phi\phi})_{-k,\ -k}\right)^{-1}(\mathbf{U}_{\tau_0}^{\phi})_{-k}$。

  直到 $||\mathbf{U}_{\tau_0}^{\phi}||_{\infty} < \epsilon$;

  重复

    更新 $\tau_0^{-1} \longleftarrow \left\{\boldsymbol{\phi}^{\mathsf{T}}\mathbf{P}_0\boldsymbol{\phi} + \mathrm{Tr}\left((-\mathbf{H}_{\tau_0}^{\phi\phi})^{-1}\mathbf{P}_0\right)\right\}/\rho(\mathbf{P}_0)$

    使用（20）更新 $\mathbf{H}_{\tau_0}^{\phi\phi}$。

  直到 收敛;

直到收敛;

算法 3： 拟合具有时变协变量的双加性治愈模型（TVcure）

目标： 拟合第 3 节中描述的 TVcure模型。

输入： 数据 $\mathcal{D} = \cup_{i=1}^{n} \{(d_{it},\ \mathbf{v}_i(t)\ \tilde{}\mathbf{v}_i(t)):\ t = 1,\ \ldots\ldots,\ t_i\}$ 以及一个 TVcure 模型设定。

输出： $\boldsymbol{\zeta} = (\boldsymbol{\psi}^\top,\ \tilde{\boldsymbol{\psi}}^\top)^\top$ 和 $\boldsymbol{\phi}$ 的估计值（对应所选惩罚参数）

$\boldsymbol{\tau} = (\boldsymbol{\lambda}^\top,\ \tilde{\boldsymbol{\lambda}}^\top)^\top$ 和 τ_0。

重复

- 选择 τ_0 并使用算法2估计 $\boldsymbol{\phi}$ （给定 $\boldsymbol{\zeta}$ 和 $\mathcal{D}$）。
- 选择 $\boldsymbol{\tau}$ 并使用算法 1 估计 $\boldsymbol{\zeta}$ （给定 $\boldsymbol{\phi}$ 和 $\mathcal{D}$）。

直至收敛;

4 模拟研究

我们开展了一项模拟研究，以评估第 3 节中所述算法利用右删失数据估计 TVcure模型各组成部分的能力，包括加性项和参考累积风险函数 $e^{\beta_0} F_0(t)$。结合第 5 节的应用，使用对应于具有总体风险函数 (4) 的扩展促发时间模型的数据生成机制，模拟了样本量为 $n = 500$ 或 1500 的 $S = 500$ 个数据集，其中 $F_0(t) = 1 - S_0(t)$，而 $S_0(t)$ 由形状参数为 2.65、尺度参数为 133 的威布尔分布的生存函数给出。(6) 和 (7) 中的回归参数取为 $\beta_0 = 0$、$\beta_1 = -.1$、$\beta_2 = .15$ 和 $\gamma_1 = .1$、$\gamma_2 = .2$，协变量相互独立，服从伯努利分布或正态分布，$z_1,\ z_3 \sim .5\,\text{Bern}(.5)$，$z_2,\ z_4 \sim \mathcal{N}(0,\ 1)$。考虑了以下加性项，

$$f_1(x_1) = -1.14 + 2.4x_1 - .88x_1^2\ ;\ f_2(x_2) = -.3\cos(2\pi x_2)$$
$$\tilde{f}_1(x_1) = .15 - .5\cos(\pi(x_1 - .75))\ ;\ \tilde{f}_2(x_3) = .6(x_3 - .5)$$

其中，$x_1 \sim$ Uniform(0, 1.5)由长期和短期生存子模型共享，而x_2、$x_3 \sim$Uniform(0, 1)是分别专属于前述各子模型的独立协变量。此类数据可根据 Yakovlev 和 Tsodikov（1996）提出的促发时间模型的生物学动机生成，其中 $(n_i|\mathbf{v}_i) \sim \text{Pois}(\vartheta(\mathbf{v}_i))$，$y_i = +\infty$if $n_i = 0$（此时个体 i“已治愈”），而 $y_i = \min\{\breve{y}_{im}:\ m = 1,\ \ldots\ldots,\ n_i\}$ otherwise，其中 $(\breve{y}_{im}|\tilde{}\mathbf{v})$ 是具有累积分布函数 $F(\cdot|\tilde{}\mathbf{v})$ 的独立同分布随机变量。独立删失时间采用 (120, 299) 上的均匀分布（情景 1）或 (60, 299) 上的均匀分布（情景 2）生成，从而得到接近 44% 或 51% 的观测（较高）右删失率，其中包括 40% 的边际（未知）治愈率。若 y_i 表示为个体 i（$i = 1, \ldots, n$）生成的未删失响应，则观测响应为 (t_i, δ_i)，其中 $t_i = \min\{y_i, c_i\}$，而 $\delta_i = I(y_i < c_i)$ 为事件指示变量。随后，第二步将这些数据转换为人月格式，见第 3.2 节；对于在随访 t 个月后仍处于风险中的个体 i （$t = 1, \ldots, 299$），其 $dt = 1$（月）以及事件指示变量为 d_{it} 。

表 1：模拟研究（情景 2）：长期和短期子模型中加性项的估计。平均绝对偏差、均方积分误差平方根（RMISE）以及 95%逐点可信区间的平均有效覆盖率。

删失	n		$f_1(x)$	$f_2(x)$	$\tilde{f}_1(x)$	$\tilde{f}_2(x)$
场景 1	500	平均绝对偏差	0.097	0.087	0.131	0.068
		均方积分误差平方根	0.127	0.112	0.174	0.097
		覆盖率	81.2	90.7	87.3	96.0
	1500	平均绝对偏差	0.053	0.053	0.070	0.042
		均方积分误差平方根	0.071	0.069	0.092	0.057
		覆盖率	93.1	93.3	94.5	96.0
情景 2	500	平均绝对偏差	0.107	0.091	0.110	0.073
		均方积分误差平方根	0.139	0.118	0.186	0.105
		覆盖率	80.2	90.5	86.6	96.2
	1500	平均绝对偏差	0.057	0.055	0.075	0.040
		均方积分误差平方根	0.077	0.072	0.010	0.056
		覆盖率	93.0	93.4	94.8	96.7

在最不利的设置下（情景 2 且 $n = 500$），使用与等距结点相关联、覆盖相应协变量观测范围的 $K = 10$ （惩罚）B样条，可以在图 1 中找到 $S = 500$ 个数据集的估计加性项及其在这 S 个重复样本上的平均值。尽管样本量适中、右删失数据比例较高且模型较为复杂，加性项的形状仍得到了相当准确的估计。表 1 报告了加性项估计质量的汇总信息。不出所料，随着样本量增加以及右删失率降低，重构效果有所改善。平均绝对偏差和均方积分误差平方根（RMISE）均随样本量增加而降低，其中 RMISE 按照 $n^{-1/2}$ 的比例下降。加性项逐点 95% 可信区间的有效（平均）覆盖率接近其名义值。对于回归参数的估计也可得出类似结论，见表 2。最后， S 个大小为 $n = 500$ 的模拟数据集的标准化累积风险函数 $F_0(t)$ 估计值也可在图 2 中找到，结果同样令人满意。

5 应用：德国女性收入与生育

本节通过真实数据展示该方法的运作方式。为此，我们利用德国的登记数据，研究女性收入与首次生育行为之间的关系。这个主题是展示本文所提出方法的理想案例。首先，德国是欧洲终身未育女性比例最高的国家之一（Kreyenfeld 和 Konietzka，2017年）。因此，每个队列中始终都有较高的“治愈比例”。其次，德国近年来实施了几项重大家庭政策

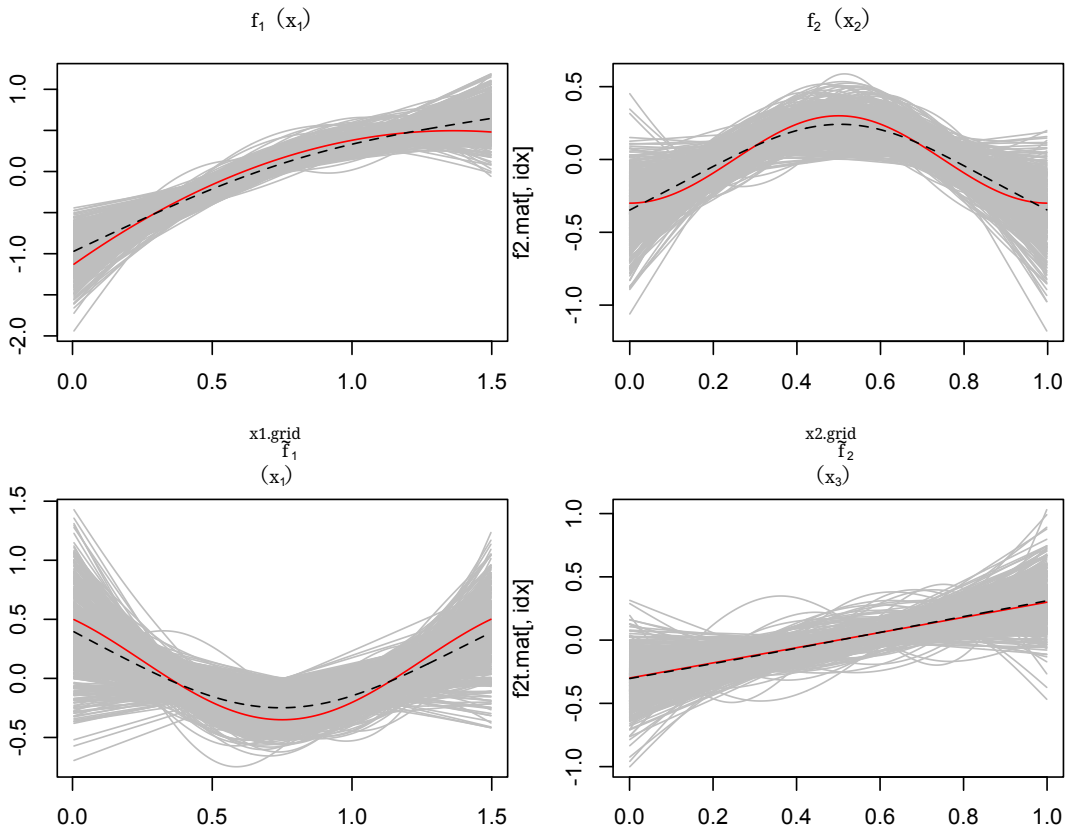


图1：模拟研究（$n = 500$ - 情景 2）：各个数据集的估计加性项（灰色）及其在 S 个重复样本上的平均值（虚线），实线（红色）对应“真实”的加性函数。

改革。其中一项非常重要的改革是2007年实施的育儿假津贴改革。与此前的规定相比，这项改革为人们在生育子女之前先进入劳动力市场提供了更强的激励。因此，人们可能会预期，女性收入与生育之间的关联在较年轻队列中比在较年长队列中更强。

所关注的数据是从2017年和2019年的德国养老金登记册中随机抽取的样本（SUFVSKT 2017/SUFVSKT 2019 数据摘录）。德国养老金登记册涵盖德国约90%的常住人口。某些职业群体（农民、律师）和公务员未纳入数据。后续分析中，我们将研究范围限定为1950—1974年出生队列的西德女性。本研究排除东德，因为德国统一之前，尤其是统一之前，两地的生育模式差异较大。此外，我们关注的是20岁时尚未生育的女性。因此，本研究不涉及青少年生育。样本包括15,248名女性，且

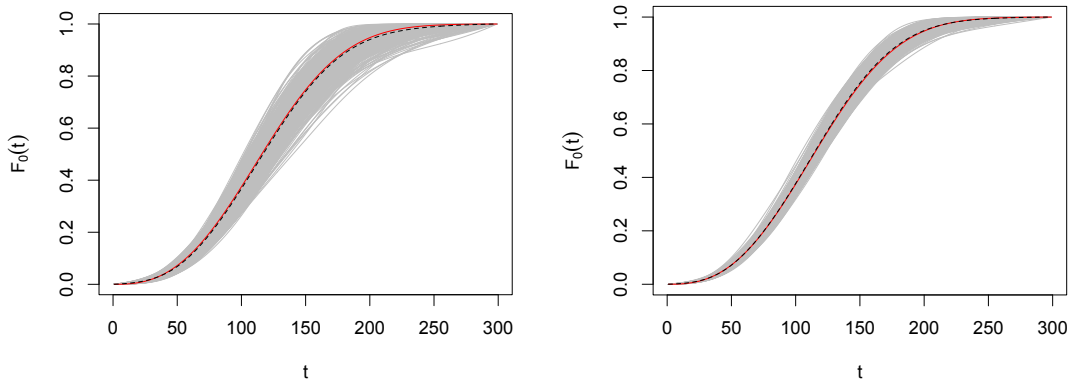


图2：模拟研究（情景 2）：每个 $S = 500$ 数据集及其平均值（虚线）在S 个重复样本中估计的累积基线风险函数$F_0(t)$，实线（红色）对应“真实”的 $F_0(t)$（左图：$n = 500$；右图：$n = 1500$）。

11,019 次首次生育。

我们关注的结果变量——妊娠开始月份，是通过第一个孩子的出生日期减去9个月计算得出的。对每位女性的随访从20岁开始，最晚至45岁；随访可能在首次妊娠时中断，也可能因失访而中断（即右删失）。若因失访而中断，则该个体最终是否“治愈”（即终身未育）存在不确定性。因此，在这一生育研究背景下，若女性到45岁时仍未生育，则视为“治愈”。数据集还包含女性完整的月度就业和收入历史。收入以收入点数记录，其中一个收入点数代表某一年的平均年收入。请注意，我们只能观察到正规就业期间的收入。正规就业是指能够获得德国养老保险基金养老金缴费积分的就业。没有正规就业收入的女性在分析中记为收入“零”。这些女性可能正在求学、失业，或因其他原因退出劳动力市场。在分析的主要部分中，我们将收入作为连续协变量。我们还通过一组随时间变化的“就业变量”控制女性是否正在求学、就业、登记失业，或因其他原因未从事正规就业。我们也控制出生队列，并将队列划分为1950—1954年、1955—1959年、1960—1964年、1965—1969年和1970—1974年出生。表3按出生队列组列出了描述性统计。此处所述的“右删失”女性，在45岁之前、首次可能妊娠发生前，随访已中断，因此其45岁时的母亲身份存在不确定性。表中显示，西德女性到45岁仍未生育的比例呈上升趋势，从1950—1955年出生队列的约21%升至1970—1974年出生队列的28%。然而，由于完整随访的比例不断下降，仍存在不确定性。

作为稳健性检验，我们还进行了额外分析，将收入分为四个类别（0, .33]（“低”），（.33, .66]（“中”），（.66,1.0）（“高”），以及

表 2：模拟研究（情景 2）：长期和短期子模型中回归参数的估计。95% 可信区间的偏倚、均方根误差和有效覆盖率。

删失	n		β_0 0.000	β_1 -0.100	β_2 0.150	γ_1 0.100	γ_2 0.200
场景 1	500	Bias	0.007	-0.006	0.010	0.007	0.005
		RMSE	0.102	0.065	0.064	0.083	0.078
		覆盖率	89.4	92.8	93.8	81.6	95.8
	1500	Bias	-0.008	0.000	0.001	0.000	0.003
		RMSE	0.053	0.035	0.037	0.045	0.045
		覆盖率	93.2	94.4	93.8	85.4	96.0
情景 2	500	Bias	0.007	-0.008	0.006	0.016	0.002
		RMSE	0.110	0.067	0.068	0.089	0.085
		覆盖率	90.4	94.4	94.2	83.8	95.4
	1500	Bias	-0.011	0.002	0.000	0.002	0.005
		RMSE	0.061	0.038	0.037	0.046	0.051
		覆盖率	91.6	94.4	95.4	85.4	93.8

表3：西德女性队列随访情况及其20岁生日与首次妊娠之间月数的汇总统计（来源：SUFVSKT 2018）。

		个体	45岁前成为母亲		
队列	n	月	Yes	No	右删失
1950—1954年	2423	277668	1900（78.4%）	519（21.4%）	4（0.2%）
1955—1959年	2388	313692	1762 (73.8%)	613 (25.7%)	13 (0.5%)
1960—1964年	3029	431144	2186 (72.2%)	797（26.3%）	46（1.5%）
1965—1969年	3385	511119	2410（71.2%）	887（26.2%）	88（2.6%）
1970—1974年	4023	655489	2761（68.6%）	1130（28.1%）	132（3.3%）

(1。00， 1.5] (‘最高’)。由于收入以收入点数衡量，切分点代表收入低于平均收入33%、处于33%至66%之间以及处于66%至平均收入之间的女性。最后一类包括收入高于平均收入的女性。图3展示了分类结果变量的分布。正如预期，较年轻女性（20—24岁）中有很大一部分正在接受教育。值得注意的是，没有从正规就业中获得任何收入的女性比例相对较高。此外，在30岁及以上的女性中，只有25%收入高于平均水平。

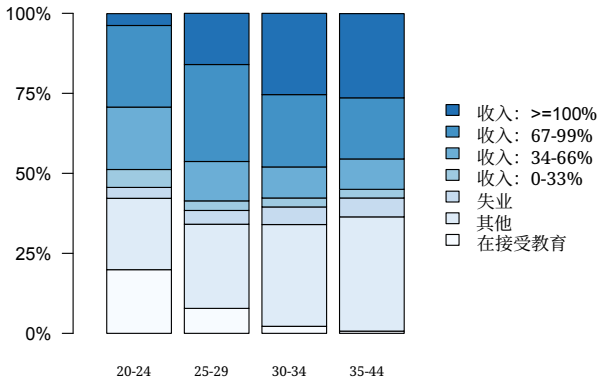


图3：西德1950—1974年出生队列未生育女性的收入和就业状态。

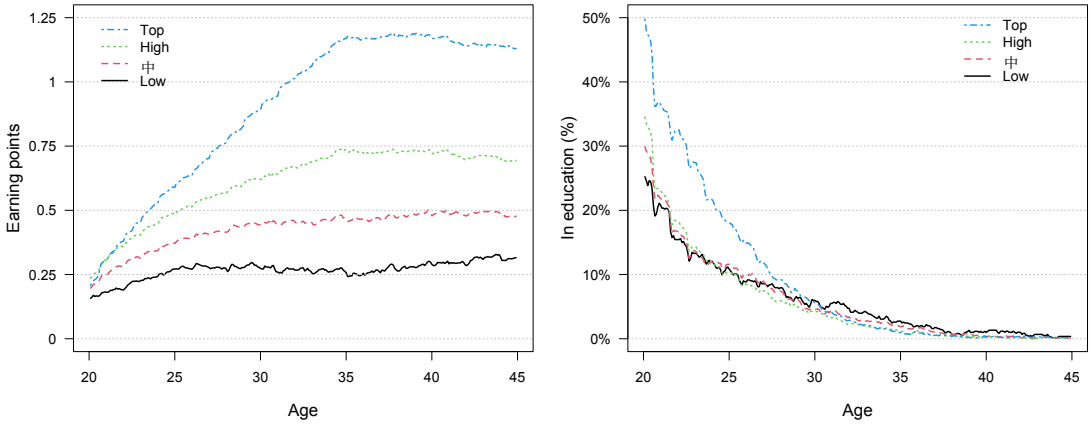


图4：西德1950—1974年出生队列未生育女性按35—44岁时最高收入划分的年龄—收入轨迹。

显然，女性的收入轨迹因职业轨道不同而大相径庭。为展示女性收入概况的差异，图4绘出了平均收入（左图）以及女性接受教育的比例（右图）

表4：拟合西德女性队列数据的TVcure模型：Status指表示就业状态的分类协变量，s(Earnings)表示对长期（Quantum）和/或短期（Timing）生存的收入加性效应。

子模型协变量		模型						
		$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$	$\mathcal{M}_5$	$\mathcal{M}_6$	$\mathcal{M}_7$
Quantum	就业状态	✓	✓	✓	✓	✓	–	–
	s(收入)	✓	✓	–	–	–	–	–
时机	就业状态	✓	✓	✓	✓	–	✓	–
	s(收入)	✓	–	✓	–	–	–	–

并按她们在35至44岁时的最高收入进行划分。图中显示，最终达到高收入水平的女性，其年龄—收入曲线在较年轻时较为平缓，但在较高年龄段则相当陡峭。较高年龄段的收入差异很大程度上可归因于教育参与的差异（右图）。这一点很重要，因为下文将使用月收入和就业状态的典型轨迹来说明模型结果。

根据第 3 节所述的 TVcure模型进行的分析，分别针对上述出生队列组开展。对数据拟合了七个 TVcure模型；长期（事件数量）和短期（时机）生存加性子模型中纳入的协变量见表4。例如，模型 $\mathcal{M}_1$ 假设就业状态和收入对事件数量和时机均有共同影响。模型拟合优度以偏差衡量，即 $D = 2\sum_{i=1}^{n}\sum_{j=1}^{n_i}\{d_{ij}\log(d_{ij}/\mu_{ij}) - (d_{ij} - \mu_{ij})\}$；模型复杂度则以有效自由度（EDF）量化，该自由度由$\mathbf{H}_\tau^{-1}\mathbf{H}_0$ 的迹给出（Hastie和Tibshirani，1990）。二者结合可得到赤池信息准则（AIC） $= D + 2$EDF。基于AIC进行的模型选择（见表5）表明，对于所有队列，就业状态（工作、学习、失业或因其他原因不在劳动力市场）会显著影响生育第一个孩子的概率以及妊娠时机。女性收入与首次生育时机和事件数量之间也存在统计学上的显著关联，但1960—1964年队列除外。对于后者，收入不会影响生育进度。图5分别绘出了各队列模型 $\mathcal{M}_1$ 的估计结果。第一行展示了一名税前收入为平均水平一半的女性的首次生育模式。图中清楚显示，生育时间表随出生队列而发生变化。首次生育有所推迟，同时分布范围扩大，这表明初为父母的年龄异质性增加。收入对事件数量和时机的加性项估计值分别显示在图5的第2列和第3列。结果表明，在20世纪50年代，收入较低的在职西德女性比收入较高者更可能在较年轻时成为母亲。不同队列间的这种关联发生了反转，其中1960—1964年队列起着关键作用。对于较近的队列，这种关联如今呈正向，收入较低会降低首次生育的可能性。关于这些模式的

表5：表4中TVcure模型的偏差、有效自由度（EDF）和赤池信息准则（AIC）

	1950—1954年出生队列			1955—1959年队列		
模型	偏差	EDF	AIC	偏差	EDF	AIC
$\mathcal{M}_1$	18055.72	14.2	<u>18084.19</u>	17460.17	14.8	<u>17489.84</u>
$\mathcal{M}_2$	18069.87	11.0	18091.81	17469.91	11.1	17492.07
$\mathcal{M}_3$	18082.72	10.9	18104.45	17491.76	11.0	17513.82
$\mathcal{M}_4$	18125.19	7.0	18139.19	17518.86	7.0	17532.86
$\mathcal{M}_5$	18134.80	4.0	18142.80	17528.57	4.0	17536.57
$\mathcal{M}_6$	18145.99	4.0	18153.99	17581.49	4.0	17589.49
$\mathcal{M}_7$	18209.75	1.0	18211.75	17644.05	1.0	17646.05

	1960—1964年队列			1965—1969年队列		
模型	偏差	EDF	AIC	偏差	EDF	AIC
$\mathcal{M}_1$	22247.38	15.4	22278.24	25007.62	15.7	<u>25038.94</u>
$\mathcal{M}_2$	22250.92	11.4	<u>22273.80</u>	25019.54	11.7	25042.88
$\mathcal{M}_3$	22263.40	11.3	22285.94	25017.12	11.3	25039.80
$\mathcal{M}_4$	22274.21	7.0	22288.21	25034.64	7.0	25048.64
$\mathcal{M}_5$	22288.15	4.0	22296.15	25045.19	4.0	25053.19
$\mathcal{M}_6$	22347.23	4.0	22355.23	25093.10	4.0	25101.10
$\mathcal{M}_7$	22491.47	1.0	22493.47	25288.45	1.0	25290.45

	1970—1974年出生队列		
模型	偏差	EDF	AIC
$\mathcal{M}_1$	29357.81	16.4	<u>29390.52</u>
$\mathcal{M}_2$	29374.80	11.8	29398.33
$\mathcal{M}_3$	29387.04	11.5	29410.10
$\mathcal{M}_4$	29404.60	7.0	29418.60
$\mathcal{M}_5$	29436.07	4.0	29444.07
$\mathcal{M}_6$	29468.11	4.0	29476.11
$\mathcal{M}_7$	29660.88	1.0	29662.88

较年轻的队列（1965—1969年和1970—1974年）非常相似：收入对生育第一个孩子这一决定的影响不断增强，而收入较高者会有所延迟。

我们还探索了一种基于TVcure模型的方法，该方法采用按月收入分类的版本EarnCat（分类方法见上文），用于稳健性检验。我们针对每个女性队列分别拟合了四个TVcure模型，在生育量或Timing子模型中分别设置有无EarnCat，并计算了偏差、有效自由度和赤池信息准则。鉴于分类导致的信息损失，各模型的偏差（为节省篇幅，此处未报告）不出所料地大于采用连续收入版本时得到的偏差。然而，对于所有队列，EarnCat对生育第一个孩子的概率以及妊娠Timing的显著联合效应所支持的定性结论是一致的。分析表明，近期队列生育子女的经济前提已经发生了变化。对于较早的队列，低收入时期会提高“生育量”；而对于较年轻的队列，情况则恰恰相反。

到目前为止，分析重点关注收入与首次生育行为之间的关联。我们尚未考虑到，一些女性年轻时收入可能较低，但人生后期收入可能较高。为此，我们利用拟合的TVCure模型，估计了四种“典型”收入轨迹下怀孕的条件概率。我们选取了以下四种情景，其描述性名称依据35—44岁时的最高收入水平而定（另见图6左上图）：

1. “低”：一名收入轨迹较低、学习至20岁的女性（实线）；2. “中”：一名学习至22岁后处于中等收入轨迹的女性（虚线）；3. “高”：一名25岁完成学业后转入高收入轨迹的女性（点线）；4. “最高”：一名27岁完成学业后转入最高收入轨迹的女性（点划线）。

我们分别按出生队列和收入组展示了估计结果（见图6上、下两幅图）。图中显示，女性收入与首次生育进程之间的关联如何随出生队列而变化。较早出生的队列中，不同女性收入概况对应的生育差异很大，但随着时间推移，这些模式变得越来越相似。我们还将图形的呈现方式倒转，展示了各收入组内部不同队列的发展情况（见图6下图）。对于收入轨迹较高和最高的女性，各队列的生育时间表相对稳定；而低收入和中等收入女性则越来越推迟首次生育。

6 讨论

所提出的方法扩展了治愈生存模型，使其能够在量子和时间加性子模型中纳入时变协变量。这些协变量并不像 Dirick et al. 的方法那样，仅限于用于非治愈个体的条件生存模型。

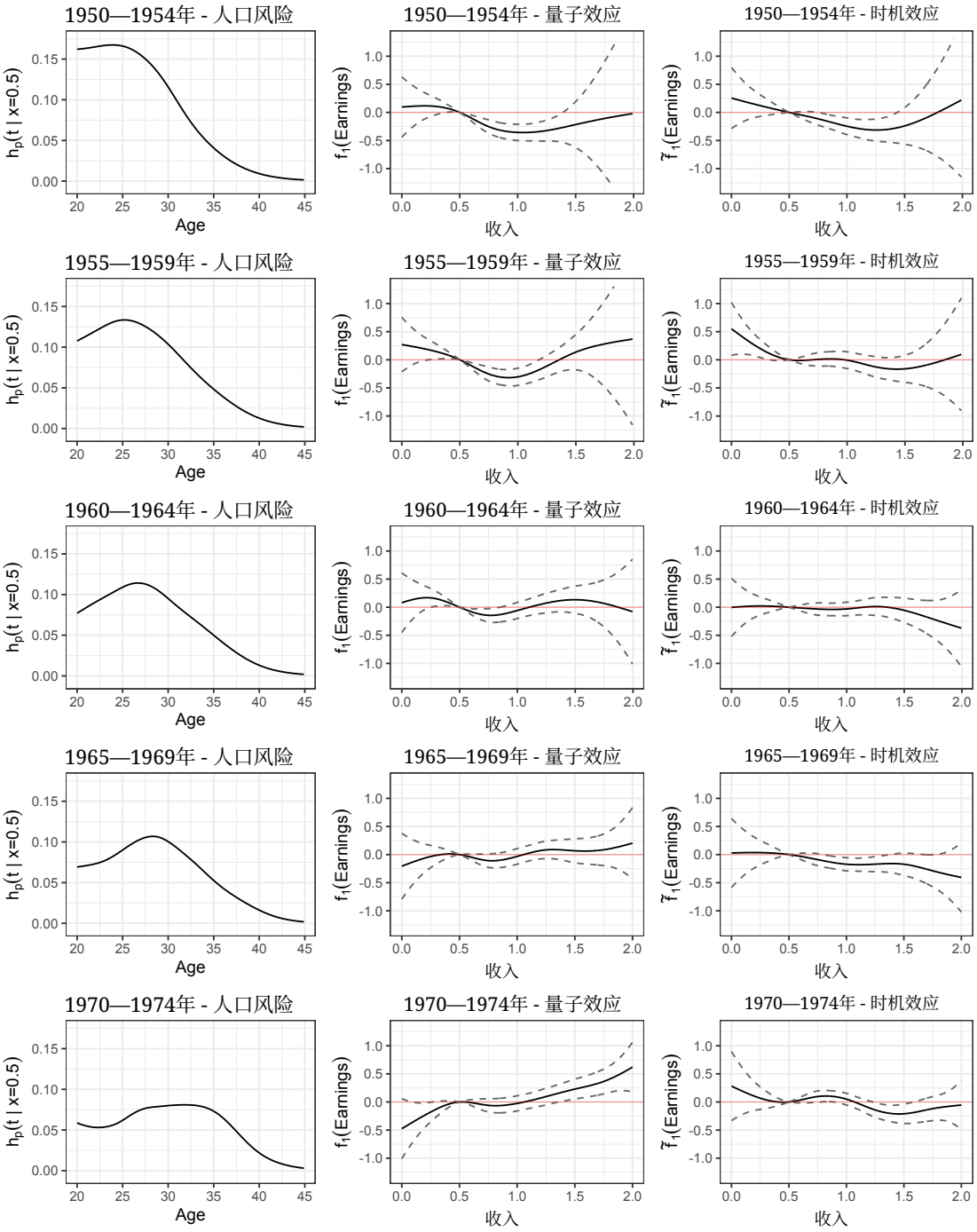


图5：模型 $\mathcal{M}_1$中收入基准值设为年度平均值的一半时，估计的基准风险率 $\mathrm{e}^{\beta_0} f_0(t)$ 以及收入对首次妊娠的 Quantum 和 Timing 的影响。

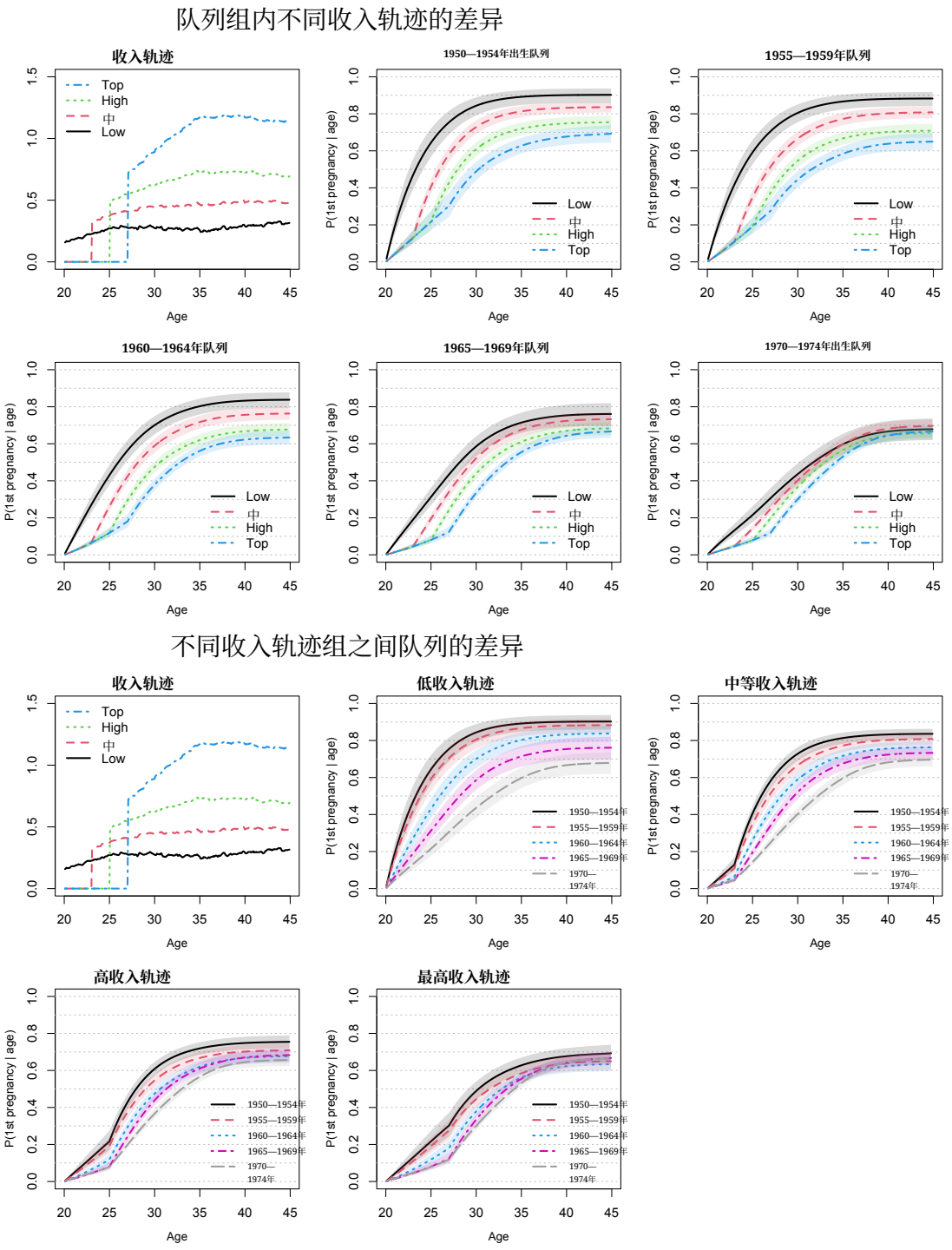


图6：按典型收入概况和出生队列划分的西德女性首次妊娠概率估计值。

（2019年）也没有像 Lambert and Bremhorst（2020年）那样限制协变量值的变化次数；后者只允许少量纯分类协变量更新。我们的新设定支持纳入时变定量解释变量，例如生育应用中频繁变动的月收入。对于这些变量的效应，也可以假设为平滑非线性形式，并通过基于拉普拉斯的近似自动选择与P样条相关的惩罚参数 $\boldsymbol{\tau}$ 边际后验分布的取值 $\boldsymbol{\tau}$。

在研究的最后一步，我们使用德国的登记数据，展示该方法如何应用于真实数据。我们考察了女性收入对首次生育行为的影响是否在德国不同队列间发生变化。数据和研究问题似乎是展示该方法实际运用过程的理想案例。德国有相当高比例的女性终身无子女。因此，每个队列的治愈比例都相当可观。此外，德国在过去几十年间实施了重大政策改革。尤其是2007年实施的育儿假津贴改革，相较以往更有力地激励人们推迟生育，直至收入达到“足够”的水平。人们预期，这项改革会降低德国高学历、事业型女性的无子女比例。我们的研究区分了生育时间与生育数量，为这一问题提供了重要且与社会政策相关的证据。我们确实发现，高收入过去会增加老队列中的无子女比例，但在较年轻的队列中已不再存在这种关系。分析表明，女性收入与首次生育之间的关联在不同出生队列间发生了变化。低收入过去会加快较老队列的生育时间并降低其无子女比例，但在近期队列中已不再观察到相同模式。事实上，如今无论收入高低，无子女比例都非常相近。女性收入不足似乎越来越多地导致德国的生育推迟。这一模式可能归因于政策改革以及与收入挂钩的育儿假福利的引入；该福利强烈激励人们推迟为人父母，直到积累足以获得适当育儿假福利的收入。

致谢

第一作者感谢 ARC项目IMAL（资助编号 20/25-107）提供的支持。该项目由瓦隆-布鲁塞尔联邦资助，并由鲁汶大学学院授予。

参考文献

Berkson, J. 和 R. Gage（1952年）。癌症患者治疗后的生存曲线。《美国统计学会杂志》*47*，501–515。

Boag, J.（1949年）。癌症治疗后治愈患者比例的最大似然估计。《皇家统计学会杂志》*B*辑（方法学）*11*(1)，11–53。

Bremhorst, V.、M. Kreyenfeld 和 P. Lambert（2016年）。德国的生育进程：使用灵活的非参数治愈生存模型进行分析。《人口学研究》*35*，505–534。Bremhorst, V. 和 P. Lambert（2016年）。使用贝叶斯P样条进行治愈生存模型的灵活估计。《计算统计学与数据分析》*93*，270–284。Chen, M.-H.、J. Ibrahim 和 D. Sinha（1999年）。一种带有存活比例的新贝叶斯模型。《美国统计学会杂志》*94*(447)，909–919。Dirick, L.、T. Bellotti、G. Claeskens 和 B. Baesens（2019年）。信用风险计算中的宏观经济因素：将时变协变量纳入混合治愈模型。《商业与经济统计学杂志》*37*，40–53。Eilers, P. H. C. 和 B. D. Marx（1996年）。利用B样条和惩罚项进行灵活平滑。《统计科学》*11*，89–102。Fahrmeir, L. 和 S. Lang（2001年）。基于马尔可夫随机场先验的广义加性混合模型贝叶斯推断。《皇家统计学会杂志》*C*辑（应用统计）*50*，201–220。Hastie, T. J. 和 R. J. Tibshirani（1990年）。《广义加性模型》。伦敦：Chapman & Hall。Jullion, A. 和 P. Lambert（2007年）。空间自适应贝叶斯P样条模型中粗糙度惩罚先验分布的稳健设定。《计算统计学与数据分析》*51*(5)，2542–2558。Kreyenfeld, M. 和 D. Konietzka（2017年）。分析无子女状态。载于 M. Kreyenfeld 和 D. Konietzka（编），《欧洲的无子女状态：背景、原因与后果》，第3–15页。施普林格国际出版公司。Lambert, P.（2021年）。利用拉普拉斯近似对右删失和区间删失数据中的非参数双加性位置-尺度模型进行快速贝叶斯推断。《计算统计学与数据分析》*161*，107250。Lambert, P. 和 V. Bremhorst（2019年）。当相同协变量影响长期和短期生存时，促发时间治愈模型中的估计与可识别性问题。《生物统计学杂志》*61*(2)，275–289。Lambert, P. 和 V. Bremhorst（2020年）。将时变协变量纳入治愈生存模型及其在生育研究中的应用。《皇家统计学会杂志》*A*辑183，333–354。Lang, S. 和 A. Brezger（2004年）。贝叶斯P样条。《计算与图形统计学杂志》*13*，183–212。

Lindsey, J. K.（1995年）。使用对数线性模型拟合参数计数过程。《皇家统计学会杂志》。*C*辑（应用统计）*44*, 201–212。

Rue, H. 和 L. Held（2005年）。高斯马尔可夫随机场：理论与应用（统计学与应用概率专著）。Chapman & Hall/CRC。

Tsodikov, A.（1998年）。一种考虑长期存活者的比例风险模型。《生物统计学》*54*, 1508–1516。

Yakovlev, A. 和 A. Tsodikov（1996年）。肿瘤潜伏期随机模型及其生物统计学应用。World Scientific Publishing Singapore。

附录

$\mathbf{U}_\lambda$ 和$\mathbf{H}_\lambda$的封闭形式表达式

令 $b_{tk} = b_k(t)$、 $\pi_t = f_0(t)\,dt = \exp\left(\sum_{k=1}^K b_{tk}\phi_k\right) / \sum_{s=1}^T \exp\left(\sum_{k=1}^K b_{sk}\phi_k\right)$, $h_{it} = h_p(t|\mathbf{v}_i(t)\ \tilde{}\mathbf{v}_i(t))$、 $F_0(t) = \sum_{s\le t}\pi_t$ 以及 $S_0(t) = 1 - F_0(t)$。对于涉及 $f_0(t)$ 的样条参数，其条件得分和精度矩阵为

$$\begin{aligned}
\mathbf{U}_{\tau_0}^{\phi} &= \frac{\partial \ell_p}{\partial \boldsymbol{\phi}} = \sum_{i=1}^{n}\sum_{t=1}^{t_i}(d_{it} - \mu_{it})\frac{\partial \log h_{it}}{\partial \boldsymbol{\phi}} - \tau_0 \mathbf{P}\boldsymbol{\phi} \\
-\mathbf{H}_{\tau_0}^{\phi\phi} &= \sum_{i=1}^{n}\sum_{t=1}^{t_i}\mu_{it}\frac{\partial \log h_{it}}{\partial \boldsymbol{\phi}}\frac{\partial \log h_{it}}{\partial \boldsymbol{\phi}^\intercal} + \tau_0\mathbf{P}
\end{aligned} \tag{20}$$

with

$$\begin{aligned}
\frac{\partial \log h_{it}}{\partial \phi_k} &= \frac{\partial \log f_0(t)}{\phi_k} + (\mathrm{e}^{\eta_F(\tilde{\mathbf{v}}_i(t))} - 1)\frac{\partial \log S_0(t)}{\phi_k} \\
\frac{\partial \log f_0(t)}{\phi_k} &= \breve{b}_{tk} = b_{tk} - \sum_s \pi_s b_{sk} \\
\frac{\partial \log S_0(t)}{\phi_k} &= -\frac{1}{S_0(t)}\sum_{s\le t}\pi_s \breve{b}_{sk}\ .
\end{aligned}$$

对于定义长期生存的回归参数和样条参数，条件得分和精度矩阵为

$$\begin{aligned}
\mathbf{U}_{\lambda}^{\psi} &= \frac{\partial \ell_p}{\partial \boldsymbol{\psi}} = \sum_{i=1}^{n}\sum_{t=1}^{t_i}(d_{it} - \mu_{it})\frac{\partial \log h_{it}}{\partial \boldsymbol{\psi}} - \mathbf{K}_\lambda(\boldsymbol{\psi} - \mathbf{b}) \\
-\mathbf{H}_{\lambda}^{\psi\psi} &= \sum_{i=1}^{n}\sum_{t=1}^{t_i}\mu_{it}\frac{\partial \log h_{it}}{\partial \boldsymbol{\psi}}\frac{\partial \log h_{it}}{\partial \boldsymbol{\psi}^\intercal} + \mathbf{K}_\lambda
\end{aligned} \tag{21}$$

其中 $\partial \log h_{it}/\partial \psi_k = (\boldsymbol{\mathcal{X}}_t)_{ik}$。对于定义短期生存的回归参数和样条参数，条件得分和精度矩阵为

$$\begin{aligned} \mathbf{U}_{\tilde{\lambda}}^{\tilde{\psi}} &= \frac{\partial \ell_p}{\partial \tilde{\boldsymbol{\psi}}} = \sum_{i=1}^{n}\sum_{t=1}^{t_i}(d_{it}-\mu_{it})\frac{\partial \log h_{it}}{\partial \tilde{\boldsymbol{\psi}}} - \tilde{\mathbf{K}}_{\tilde{\lambda}}(\tilde{\boldsymbol{\psi}}-\mathbf{g}) \\ -\mathbf{H}_{\tilde{\lambda}}^{\tilde{\psi}\tilde{\psi}} &= \sum_{i=1}^{n}\sum_{t=1}^{t_i}\mu_{it}\frac{\partial \log h_{it}}{\partial \tilde{\boldsymbol{\psi}}}\frac{\partial \log h_{it}}{\partial \tilde{\boldsymbol{\psi}}^{\mathsf{T}}} + \tilde{\mathbf{K}}_{\tilde{\lambda}} \end{aligned} \tag{22}$$

其中 $\partial \log\ h_{it}/\partial \tilde{\psi}_k = (\ \tilde{}\boldsymbol{\mathcal{X}}_t)_{ik}\big(1+\mathrm{e}^{\eta_F(\tilde{}\mathbf{v}(t))}\log\ S_0(t)\big)$。交叉导数的闭式表达式（与惩罚参数无关）也可以得到：

$$-\mathbf{H}^{\psi\tilde{\psi}} = \sum_{i=1}^{n}\sum_{t=1}^{t_i}\mu_{it}\frac{\partial \log h_{it}}{\partial \boldsymbol{\psi}}\frac{\partial \log h_{it}}{\partial \tilde{\boldsymbol{\psi}}^{\mathsf{T}}}\,.$$

令 $\boldsymbol{\zeta} = (\boldsymbol{\psi}^{\mathsf{T}},\ \tilde{\boldsymbol{\psi}}^{\mathsf{T}})^{\mathsf{T}}$ 和 $\boldsymbol{\tau} = (\lambda^{\mathsf{T}},\ \tilde{\lambda}^{\mathsf{T}})^{\mathsf{T}}$。则，$\boldsymbol{\zeta}$ 的得分函数和（负的）精度矩阵为

$$\mathbf{U}_{\tau}^{\zeta} = \begin{pmatrix} \mathbf{U}_{\lambda}^{\psi} \\ \mathbf{U}_{\tilde{\lambda}}^{\tilde{\psi}} \end{pmatrix} \quad ; \quad \mathcal{H}_{\tau} = \begin{bmatrix} \mathbf{H}_{\lambda}^{\psi\psi} & \mathbf{H}^{\psi\tilde{\psi}} \\ \mathbf{H}^{\tilde{\psi}\psi} & \mathbf{H}_{\tilde{\lambda}}^{\tilde{\psi}\tilde{\psi}} \end{bmatrix}\,. \tag{23}$$