用于相干分析和预测死因别死亡率的狄利克雷-多项式-泊松框架

Andrea Nigri福贾大学社会科学系

Han Lin Shang 精算研究与商业分析系 麦考瑞大学

Francesco Ungolo 风险与精算研究学院 新南威尔士大学悉尼分校

arXiv:2603.00973v1 [stat.AP] 1 Mar 2026

摘要

分别对死因特异性死亡率及其预测进行建模，可能会产生不一致的预测，因为各死因死亡人数之和可能与观测到的总死亡人数不符。我们针对死因死亡人数构建了一个层级概率框架，其中死亡总数和各死因死亡的发生情况均视为随机变量。在给定死亡总数的条件下，各死因死亡人数服从多项分布，而总人数则采用泊松分布建模，死因概率向量则服从狄利克雷分布。泊松模型和狄利克雷模型均捕捉了死因特异性死亡率随年龄和日历年的变化，从而在保留易于解释的人口统计模式的同时，从构造上确保相干性。这种模型构建方式自然地保持了各死因死亡人数之和与总死亡率之间的相干性。

该方法通过分析美国和法国的死因特异性死亡率来展示，数据来源于人类死亡率数据库，涵盖1979年至2023年，并按性别和年龄分别分析，同时将死亡按主要死因类别分组。实证分析采用滚动 15 年样本外评估，并将所提出的模型与标准 Lee–Carter模型及其组成数据扩展进行比较。结果表明，该方法能够对不同国家和性别进行协调一致的预测，达到有竞争力的预测准确度，并且对总体死亡率和死因特异性死亡率的不确定性估计均经过良好校准。

关键词：年龄-时期模型；贝叶斯层级建模；组成数据计数；狄利克雷-多项式-泊松模型；李-卡特模型

1 引言

由于生活水平的提高、医学进步以及公共卫生政策的发展，二十世纪期间，发达国家的人类预期寿命显著增加。例如，在美国（美国），出生时预期寿命从1900年的47.3岁增加到2010年的78.7岁（Crimmins，2015）。这一现象带来了深刻的流行病学和社会挑战。除了衰老过程之外，各种死因的影响也会随时间发生变化。由于医学进步，某些死因（如传染病）导致的死亡比例正在大幅下降，而另一些死因则持续存在，甚至有所上升。对于公共卫生决策者而言，识别按死因划分的死亡率演变，对于指导预防策略、分配资源以及设计针对特定队列或世代的干预措施具有重要意义（例如参见 Preston 等人，2000年；Booth 和 Tickle，2008；Basellini 等人，2023年）。

传统的总体死亡率外推方法，例如 Lee 和 Carter (参见 Basellini 等人，2023年的最新综述），已被证明非常适合捕捉年龄特定趋势并预测总体死亡率。然而，在对亚群体或按死因划分的死亡率进行单独分析时，这些模型可能会产生违反基本人口学约束的预测，例如各死因特异性死亡人数之和应等于总死亡人数这一约束（例如参见 Wilmoth，1995年；Hyndman 等人，2013）。这种不相干性会导致不合理的长期预测，并使其难以用于政策目的。

从统计学视角来看，按死因对死亡率进行建模绝非易事。顾名思义，各死因互斥且穷尽所有情况。其构成结构意味着，像 Chiang (1991) 的开创性研究那样，对每种死因进行独立建模，可能会产生不相干的结果，即各死因预测值之和超过或低于总死亡率。这种不相干性损害了精算应用和长期公共卫生情景的可信度，且往往会导致更悲观的预测（Wilmoth，1995年）。越来越多的研究方向开始对死因事件比例进行组成数据统计分析，并采用能够确保相干性的统计框架（例如参见Oeppen，2008；Kjærgaard 等人，2019年；Cardillo 等人，2023年，2024等）。

这一方向的一项有前景的进展，是将狄利克雷回归纳入特定年龄的死亡率建模。通过假设按死因划分的死亡比例服从狄利克雷

分布，便能直接考虑死亡率的组成数据特性，而无需像郑和陈(2017年)的研究那样进行转换。近期研究表明，这种方法在保持可解释性的同时，还提供了一个连贯的框架，用于捕捉不同死因之间的年龄和时间变化（Graziani和Nigri，2026年）。此外，狄利克雷框架可以自然地与年龄–时期（AP）模型及其年龄–时期–队列扩展相结合。

年龄–时期–队列模型的分解揭示了衰老的生物学影响、诸如大流行病或医疗创新等时期冲击，以及反映不同世代独特经历的队列效应（Renshaw 和 Haberman，2006年；Hunt 和 Blake，2014）。将这一范式扩展至死因研究，有助于研究人员发现死亡率动态中隐藏的规律，并改进整体预测。

类似研究主要关注死亡率差异的年龄–时期模型，尤其是比较癌症和循环系统疾病等主要死因时。利用斯凯拉姆分布等概率模型，这些研究能够以较为简约的方式刻画死亡人数的差异，凸显年龄–时期结构在分析死因特异性差异方面的潜力（李等，2019年；Lanfiuti Baldi 等，2025年）。这些研究共同凸显了一项关键的方法学挑战：如何在保留灵活性以刻画年龄和时期特异性动态的同时，对总死亡率水平以及按死因划分的死亡构成进行建模。

本文提出狄利克雷–多项式–泊松（DMP）框架来分析分死因死亡率，为相关文献作出了贡献。该框架假设各死因的死亡人数服从多项分布，其中死因特异性死亡概率和死亡总数均为随机变量，分别服从狄利克雷分布和泊松分布。该模型带来双重离散度来源，能够捕捉死亡人数以及各死因发生比例的变化。这两类变化均可由年龄和日历年等协变量解释。本文展示了该方法在分析和预测美国与法国1979年至2023年观测到的分死因死亡率中的应用。此类死亡率按性别分别分析，并采用滚动留出法，以观测期最后15年的死亡率评估两种方法的预测能力。本文将DMP框架的结果与用于分析各死因死亡率的李–卡特

（LC）模型，以及Kjærgaard 等人（2019年）的组成数据分析（CoDa）方法进行比较。

DMP下的模型参数通过完全贝叶斯分析进行估计。这一推断策略使我们能够量化预测率的不确定性。这一方面对政策与规划具有重要意义，因为它使人们能够评估模型或参数的不确定性水平；更全面地比较不同方法的估计或预测准确度；并基于各种假设探索不同情景（Chatfield，2000年）。

预测不确定性由联合生成模型构造而得，并自然地整合了（i）未来死亡计数和死因分配中的随机变异，以及（ii）通过后验分布体现的参数不确定性。这不同于经典方法，例如按死因分别拟合的LC模型，或应用于生命表数量的CoDa 方法；在这些方法中，不确定性区间往往是事后引入的，可能无法以统一方式传播所有不确定性来源，也不一定能同时为总量和死因给出一致的不确定性表述。通过将人口学推理与先进的统计建模相结合，我们的方法同时满足了监测死因模式变化的流行病学需求，以及对一致且校准良好的预测的统计要求。这一贡献迈向了一个将死亡率预测与死因分析统一起来的一般统计框架，其应用范围从长寿风险管理到卫生政策规划。

本文其余部分安排如下。第 2 节介绍死亡率建模框架、DMP框架及其应用。第 3 节阐述贝叶斯推断框架。第4节介绍一种使用 AP 和 LC 模型的预测框架。第 5 节详细介绍实证分析所用的数据。第 6 节报告并讨论结果，第 7 节总结全文，并提出未来研究方向。

2 长寿建模

死亡率的统计建模捕捉了三个关键维度的变化：年龄、时期和死因。年龄描述了在整个生命周期中系统性变化的生物和社会过程；时期反映了同期冲击和创新，例如

大流行病、战争或医学进步，这些因素会在特定日历年内同时影响所有年龄组。最后，死因将死亡率划分为互斥且穷尽的流行病学类别。

为此，令 $E_{a,t}$ 表示风险暴露量（风险人年），即在时期 t内年龄组 a 的风险暴露量；令$Y_{a,t}$ 表示在时期 t 因死因 c导致的、最近一次生日年龄为 a 的个体死亡随机数。因此，年龄组a 在时期 t的死亡总数 $Y_{a,t}$，可通过对各死因的死亡数$Y_{a,t} = \sum_c Y_{a,t,c}$求和得到。

对于每个年龄 a、年份 t和死因 c，研究者会观测到以下数据：

- E_{at}：人口风险暴露量，通常近似为时期 t 开始时年龄为 a 的人口与同一时期结束时年龄为 a 的人口的平均值；

- Y_{atc}：在年份 t、年龄为 a时，因死因 c导致的观测死亡数。

相应的中心死亡率定义为

$$m_{a,t,c} = \frac{Y_{a,t,c}}{E_{a,t}}, \tag{1}$$

这表示死因 c 在年份 t、年龄 a 的平均死亡强度。在人口统计和精算分析中，该量用作死亡力 $\mu(a,t,c)$的无偏估计量，即在精确年龄 a 和日历时间 t 因死因 c (Preston 等人,2000年):

$$m_{a,t,c} \approx \mu(x,t,c), \quad x \in [a, a+1). \tag{2}$$

因此，沿袭早期文献的做法（例如，Brouhns等人，2002），假设死亡随机数 $Y_{a,\ t,\ c}$ 服从泊松分布，其均值和方差取决于中心死亡率，如下所示：

$$Y_{a,t,c} \sim \text{Poisson}(E_{a,t}\, m_{a,t,c}).$$

这一表述 e 确保死因特异性死亡数与死亡总数相协调。

根据泊松分布的可重复性，死因特异性死亡数之和也服从泊松分布：

$$\sum_{c} Y_{a,t,c} = Y_{a,t} \sim \text{Poisson}\left(E_{a,t} m_{a,t}\right), \tag{3}$$

其中$m_{a,\ t} = \sum_{c} m_{a,\ t,\ c}$。

所提出的 DMP 框架由一个层次模型构成，其中：i）死亡总数 $Y_{a,t}$ 服从泊松分布；ii）按各死因划分的死亡人数向量$Y_{a,t} = (Y_{a,t,1}, \ldots, Y_{a,t,C})$服从狄利克雷-多项式（DM）分布（Mosimann，1962年）。为便于说明，表2列出了各死因。DM 分布由两个参数确定，即死亡总数 $Y_{a,t}$ 和向量 $\alpha_{a,t} = \phi\gamma_{a,t}$，其中 $\gamma_{a,t} = (\gamma_{a,t,1}, \ldots, \gamma_{a,t,C})$，满足 $\sum \gamma_{a,t,c} = 1$，而 ϕ 控制 $\gamma_{a,t}$ 的过度离散。因此，${}_{c}Y_{a,t} \sim \text{DM}\left(\alpha_{a,t}, Y_{a,t}\right)$的概率质量函数为

$$p\left(\overline{y}_{a,t} \mid \overline{\alpha}_{at}, y_{a,t}\right) = \frac{\Gamma(\phi)\,\Gamma\left(y_{a,t} + 1\right)}{\Gamma\left(y_{a,t} + \phi\right)} \prod_{c=1}^{C} \frac{\Gamma\left(y_{a,t,c} + \phi\gamma_{a,t,c}\right)}{\Gamma\left(\phi\gamma_{a,t,c}\right)\Gamma\left(y_{a,t,c} + 1\right)},$$

其中，$\Gamma(\cdot)$ 表示伽马函数。

这种构造确保死因特异性死亡风险可分解为$m_{a,\ t,\ c} = m_{a,\ t}\gamma_{a,\ t,\ c}$，从而使狄利克雷-多项模型的均值结构与总体死亡风险$m_{a,\ t}$ 以及组成比例$\gamma_{a,\ t}$相一致。每个概率$\gamma_{a,\ t,\ c}$ 均按如下 softmax 函数进行建模：

$$\gamma_{a,t,c} = \frac{\exp\left(\eta_{a,t,c}\right)}{\sum_{c} exp\left(\eta_{a,t,c}\right)}. \tag{4}$$

其中，$\eta_{a,\ t,\ c}$ 是死因 c在年龄 a和年份 t下的潜在线性预测量。

DM 模型将多项分布与服从狄利克雷分布的概率相结合。这样，既可以处理死因特异性死亡数向量中的过度离散，也可以考虑医疗和社会进步带来的历年构成变化。此外，由于狄利克雷分布的支持集为单纯形，最终模型能够确保组成一致性，并为概率建模和预测提供灵活的工具

建模与预测。

对于第5节所述的数据，考虑两个模型来描述泊松分布的速率$m_{a,\,t}$以及 softmax 函数的参数 η_{atc} ，以推导 DM 模型的底层概率，即第2.1 节中的年龄-时期模型和第2.2节中的 LC 模型。

2.1 年龄-时期模型

该模型假设年龄和时期是两个按比例影响泊松分布速率的因素：

$$\log m_{a,t} = \nu_0 + \delta_a + \pi_t, \tag{5}$$

其中， ν_0 为基线对数死亡率， δ_a 为对应于年龄 a相关效应的因子偏差， π_t 为日历年 t带来的比例效应。类似地，对于死因 c ，年龄和时期分别通过参数$\zeta_{a,\,c}$ 和 $\lambda_{t,\,c}$对参数 $\eta_{a,\,t,\,c}$ 产生加性效应，作为相对于基线参数 $\varphi_{0,\,c}$的偏差：

$$\eta_{a,t,c} = \varphi_{0,c} + \zeta_{a,c} + \lambda_{t,c}. \tag{6}$$

死亡率通常会随年龄增长和日历年推移而平滑变化，死因别死亡概率也呈现相同的模式。为反映这一经验规律并避免过拟合，我们对参数施加平滑效应 δ_a， π_t， η_a^c，以及 $\eta_t^c(c = 1, \ldots\ldots, C)$。这一策略提高了人口稀少单元中估计值的稳定性，并捕捉死亡率数据中常见的渐进式人口和流行病学转变。为此，将每种效应建模为二阶高斯随机游走（RW2，Bryant 和 Zhang(2018)）：

$$\Delta^2\delta_a \equiv \delta_a - 2\delta_{a-1} + \delta_{a-2} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\delta^2\right), \quad a \geqslant 3 \tag{7}$$

$$\Delta^2\pi_t \equiv \pi_t - 2\pi_{t-1} + \pi_{t-2} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\pi^2\right), \quad t \geqslant 3$$

$$\Delta^2\zeta_{a,c} \equiv \zeta_{a,c} - 2\zeta_{a-1,c} + \zeta_{a-2,c} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\zeta^2\right), \quad a \geqslant 3 \tag{8}$$

$$\Delta^2\lambda_{t,c} \equiv \lambda_{t,c} - 2\lambda_{t-1,c} + \lambda_{t-2,c} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\lambda^2\right), \quad t \geqslant 3.$$

对于 $a = 1$、2 和 $t = 1$，2，年龄和时期特异性参数是自由参数，将通过第 3 节所述的贝叶斯推断程序进行推断。

2.2 李-卡特模型

该模型假设泊松分布的对数率由基准死亡率水平 ν_0、年龄特异性因子 δ_a 和年龄-时期交互效应 [1] 之和给出：

$$\log m_{a,t} = \nu_0 + \delta_a + \beta_a \kappa_t. \tag{9}$$

为确保模型参数的可识别性，施加经典李-卡特约束（Hunt 和 Blake，2014）：

$$\sum_{t=1}^{T} \kappa_t = 0, \qquad \sum_{a=1}^{A} \beta_a = 1.$$

通过假设年龄效应 δ_a 遵循二阶随机游走过程（RW2过程）并依据方程（7）进行平滑；而对于LC模型， κ_t 则假设遵循一阶随机游走（RW1）：

$$(\kappa_t \mid \kappa_{t-1}) \sim \mathcal{N}(\kappa_{t-1}, \sigma_\kappa^2), \qquad t \geqslant 2.$$

参数$\eta_{a,\ t,\ c}$ 被建模为年龄和死因-时期交互项的加性函数。更具体地说，该交互项采用了与泊松模型对数率中相同的日历年效应κ_t。

这种方法能够捕捉不同死因之间死亡总数的系统性重新分配：

$$\eta_{a,t,c} = \varphi_{0,c} + \zeta_{a,c} + \vartheta_c \kappa_t,$$

其中 $\zeta_{a,\ c}$ 是年龄-死因平滑偏差， ϑ_c 是死因特异性载荷，用于衡量死因概率与时期之间的交互作用。

死因-时期交互项使用与泊松对数率相同的日历年份效应 κ_t。在

[1]这种参数化在代数上等价于经典 LC 模型（$a_a = \nu_0 + \delta_a$， $b_a = \beta_a$）。在我们的方法中，这种重新参数化具有以下优点：将总体死亡率水平（ν_0）与年龄分布（δ_a）分离；便于对年龄效应采用中心化平滑先验；并在保持相同均值结构的同时提高马尔可夫链蒙特卡洛稳定性。通过 $\sum_a \delta_a = 0$、 $\sum_t \kappa_t = 0$ 和 $\sum_a \beta_a = 1$ 实现模型识别。

LC–DM模型中，泊松和狄利克雷-多项参数联合在单一贝叶斯层次结构中估计（即不对 κ_t采用两阶段插件法），这可由第 3 节的联合似然明确看出。

年龄别死因特异性效应$\zeta_{a,\,c}$ 也通过 RW2 过程进行平滑处理，如式（8）所示。

3 推断

3.1 先验分布

第5 节对死因特异性死亡率进行统计分析时，采用非信息先验假设。所有参数先验上均假定成对独立先验地，因此先验信息不会对模型各组成部分之间施加人为依赖关系。

因此，对控制总死亡率和死因特异性构成的截距和水平参数赋予弥散先验。AP 和 LC 设定下各模型参数的先验分布汇总于表1。

表1：先验　AP–DM 和 LC–DM 模型的规格与可识别性约束。

模型块	参数	先验 / 约束
AP–DM：总死亡率（泊松）		
	ν_0	$\mathcal{N}(0, 5^2)$
	δ_1, δ_2	$\mathcal{N}(0, 5^2)$
	$\pi_1,\ \pi_2$	$\mathcal{N}(0, 1)$
	σ_δ	$\mathcal{N}^+(0, 1^2)$
	σ_π	$\mathcal{N}^+(0, 0。3^2)$
AP–DM：构成（狄利克雷-多项），针对每个死因 $c = 1,\ \ldots\ldots,\ C$		
	$\varphi_{0,c}$	$\mathcal{N}(0, 5^2)$
	$\zeta_{1,}{}^{c,}{}_{,c}{}^{\zeta_2}$	$\mathcal{N}(0, 5^2)$

下页续

模型块	参数	先验 / 约束
	$\lambda_{1,c}, \lambda_{2,c}$	$\mathcal{N}(0,1)$
	σ_ζ	$\mathcal{N}^+(0,1^2)$
	σ_λ	$\mathcal{N}^+(0,0.3^2)$
AP–DM：过度离散		
	ϕ	$\log \mathcal{N}(\phi_\mu, \phi_\sigma)$
AP–DM：可识别性约束		
		$\sum_{a=1}^{A} \delta_a = 0,\ \sum_{t=1}^{T} \pi_t = 0,\ \sum_{c=1}^{C} \varphi_{0,c} = 0,$ $\sum_{a=1}^{A} \zeta_{a,c} = 0,\ \sum_{t=1}^{T} \lambda_{t,c} = 0\ \ (\forall c)_\circ$
LC–DM：总死亡率（LC +泊松）		
	ν_0	$\mathcal{N}(0,10^2)$
	$\delta_1,\ \delta_2$	$\mathcal{N}(0,5^2)$
	β_a	$\mathcal{N}^+(0,1^2)$
	κ_1	$\mathcal{N}(0,1^2)$
	σ_δ	$\mathcal{N}^+(0,0.5^2)$
	σ_κ	$\mathcal{N}^+(0,0.5^2)$
LC–DM：构成（狄利克雷-多项），针对每个死因 $c = 1,\ \ldots\ldots,\ C$		
	$\varphi_{0,c}$	$\mathcal{N}(0,1^2)$
	$\zeta_{1,c}, \zeta_{2,c}$	$\mathcal{N}(0,1^2)$
	ϑ_c	$\mathcal{N}(0,0.5^2)$
	σ_ζ	$\mathcal{N}^+(0,0.4^2)$
LC–DM：过度离散		
	ϕ	$\log \mathcal{N}(\phi_\mu, \phi_\sigma)$
LC–DM：可识别性约束		
		$\sum_{t=1}^{T} \kappa_t = 0,\ \sum_{a=1}^{A} \beta_a = 1,\ \sum_{a=1}^{A} \delta_a = 0,$ $\sum_{c=1}^{C} \varphi_{0,c} = 0,\ \sum_{a=1}^{A} \zeta_{a,c} = 0\ \ (\forall c)_\circ$

该模型包含一小部分参数（例如截距、离散度和分死因载荷），而其余参数按年龄和时期索引（例如 δ_{a}、 β_{a}、 κ_{t}、 $\zeta_{a,\ c}$）。因此，总维度取决于数据，并随所考虑的年龄数（A）和时期数（T）增加。

这些选择旨在尽量降低先验设定对后验推断的影响，使数据能够完全决定参数的后验分布。

3.2 似然

设 θ 表示在AP或LC模型假设下索引DMP的全部模型参数向量。对于每个年龄组 a 和日历年 t，在给定相应参数的条件下，观测到的死亡人数条件独立。

θ的似然函数在给定暴露量$E_{a,\ t}$ 以及按死因划分的观测死亡人数$\mathbf{Y}$的条件下，分解为死亡总数的泊松分量和死因构成的DM分量：

$$\mathcal{L}(\theta;\mathbf{Y},\mathbf{E})=\prod_{a=1}^{A}\prod_{t=1}^{T}\underbrace{p\left(y_{a,t}\mid E_{a,t}m_{a,t}\right)}_{\text{Poisson}\left(E_{a,t}m_{a,t}\right)}\times\underbrace{p\left(\overline{y}_{a,t}\mid\phi\overline{\gamma}_{a,t},y_{a,t}\right)}_{\text{DM}\left(\phi\overline{\gamma}_{a,t},y_{a,t}\right)},\tag{10}$$

其中$y_{a,\ t}=\left(y_{a,t,1}\cdots,y_{a,t,C}\right)$、$\overline{y}_{a,\ t}=\sum_{c=1}^{C}y_{a,\ t,\ c}$，且 θ 可以为 $\theta_{\text{AP}}=\{\nu_{0},\delta,\pi,\phi,\zeta,\lambda,\sigma_{\delta},\sigma_{\pi},\sigma_{\zeta},\sigma_{\lambda}\}$、$\theta_{\text{LC}}=\{\nu_{0},\delta,\beta,\kappa,\phi,\zeta,\vartheta,\sigma_{\delta},\sigma_{\kappa},\sigma_{\zeta}\}$。

泊松分量支配死亡总数的不确定性，而狄利克雷-多项分量刻画这些死亡在各死因之间的分配，并兼顾过度离散与构成结构。

3.3 后验推断

后验推断由第3.2 节中的似然与第3.1节所述的先验设定密度之积得出。所得后验分布不具有闭式表达式。

使用哈密顿蒙特卡洛（HMC）算法估计参数的后验分布。该算法属于 Metropolis-Hastings 采样器的一类，能够更快速、更有效地

探索参数空间。该算法通过 Stan开发团队 的接口，在 Stan（2025年）中实现，使用的程序包为 `rstan`。采样器通过 Hoffman 和 Gelman （2014）提出的“无转向采样器”自动调优。HMC采样器的一项关键特性是能够生成沿参数空间中有指导性的轨迹前进的候选样本，并利用目标密度的几何结构。这使探索更加高效，尤其是在参数高度相关的高维参数空间中，从而降低自相关并加快收敛。

4 预测

令 H 表示预测期，$\mathcal{T}^{+} = \{T+1, \ldots\ldots T+H\}$ 表示未来时期集合。我们的主要目标是获得相干的按死因划分的年龄别死亡率预测，使每个年龄和时期的各死因死亡率之和等于相应的总死亡率。这可确保流行病学可解释性，并避免汇总预测与细分预测之间出现不相干。

如果有官方人口预测数据，则可纳入暴露量 $E_{a,\,t}$，以生成未来死亡计数预测。然而，我们模型的基本输出是率尺度上的预测：总死亡率 $m_{a,\,t}$，以及死因别占比向量 $\gamma_{a,\,t,\,c}$。这种策略可确保预测在各死因之间保持内部一致，即使暴露量未知或不确定也是如此。

基于 AP–DM 的预测。 通过将假设的二阶随机游走动态延伸至 T 之后，可以获得未来时期效应。对于每个后验抽样和 $t \in \mathcal{T}^{+}$,

$$\pi_t = 2\pi_{t-1} - \pi_{t-2} + \varepsilon_t^{(\pi)}, \quad \varepsilon_t^{(\pi)} \sim \mathcal{N}(0, \sigma_\pi^2),$$
$$\lambda_{t,c} = 2\lambda_{t-1,c} - \lambda_{t-2,c} + \varepsilon_{t,c}^{(\lambda)}, \quad \varepsilon_{t,c}^{(\lambda)} \sim \mathcal{N}(0, \sigma_\lambda^2), \quad c = 1, \ldots, C.$$

年龄效应 δ_a 和 $\zeta_{a,\,c}$ 保持不变，并应用中心化约束以确保数值稳定性。因此，投影总率$m_{a,\,t}$ 以及相应的死因占比 $\gamma_{a,\,t,\,c}$ 计算如下

:

$$m_{a,t} = \exp\{\nu_0 + \delta_a + \pi_t\}, \qquad \gamma_{a,t,c} = \frac{\exp\{\varphi_{0,c} + \zeta_{a,c} + \lambda_{t,c}\}}{\sum_{c'} \exp\{\varphi_{0,c'} + \zeta_{a,c'} + \lambda_{t,c'}\}}, \qquad \sum_c \gamma_{a,t,c} = 1.$$

基于 LC–DM 进行预测。对于 LC–DM，我们通过延续 RW1 动态过程来模拟未来 κ_t：

$$\kappa_t = \kappa_{t-1} + \varepsilon_t^{(\kappa)}, \qquad \varepsilon_t^{(\kappa)} \sim \mathcal{N}(0, \sigma_\kappa^2), \qquad t \in \mathcal{T}^+.$$

年龄效应 δ_a 和 ζ_{ac} 保持不变（采用中心化约束），而 β_a 和 ϑ_c 是随时间不变的载荷。

对于 LC–DM，相应的预测量为

$$m_{at} = \exp\{\nu_0 + \delta_a + \beta_a \kappa_t\}, \qquad \gamma_{atc} = \frac{\exp\{\varphi_{0,c} + \zeta_{a,c} + \vartheta_c \kappa_t\}}{\sum_{c'} \exp\{\varphi_{0,c'} + \zeta_{a,c'} + \vartheta_{c'} \kappa_t\}}, \qquad \sum_c \gamma_{a,t,c} = 1.$$

对于 AP 和 LC 模型，这些预测量组合后可直接得到死因特异性死亡率$m_{a,\,t,\,c} = m_{a,\,t}\,\gamma_{a,\,t,\,c}$，其构造本身保证了协调一致，因为$\sum_c m_{atc} = m_{at}$对所有 (a,t)均成立。因此，预测自然同时给出总体动态及其在各死因间的分解 [2]。

5 数据

第2节概述的模型通过对死亡率的分析加以展示，死亡率数据取自人类死亡率数据库（2026年）。这些数据摘自 HMD 1979年到2023年的按性别和五岁年龄组（0–100+）编制的时期生命表。死因死亡人数和暴露量取自美国和法国的人类死因数据系列（2026年）（HCD），涵盖两种性别，并据此计算按死因、年龄和日历年划分的死亡构成。HCD 数据库提供高质量的五岁年龄组死因特异性死亡率数据。死亡原因的分类体系依据

[2]如果有暴露量 E_{at}，则可以通过以下方式模拟预测计数

$$Y_{a,t} \sim \text{Poisson}(E_{a,t} m_{a,t}), \qquad \overline{Y}_{a,t} \sim \text{DM}(\phi \overline{\gamma}_{a,t}),$$

在设计上与$\sum_c Y_{a,\,t,\,c} = Y_{a,\,t}$相一致。

世界卫生组织，并由三个汇总层级组成：完整列表、中间列表和简表。每种分类均采用各国统一的标准制定，以确保其同质性和可比性。利用这些数据，可以建立一套通用且标准化的方法，在五岁年龄组内将死亡重新分配到104个疾病类别，从而避免因ICD修订版而产生的问题，并在编码实践存在差异的情况下确保跨国可比性。从HCD分类出发，主要死因按年龄分组，直至 100+。

为了解可能影响美国人口预期寿命的情况，考虑了六大主要死因：传染性疾病、肿瘤、心血管疾病（心血管疾病）、呼吸系统疾病、外因疾病以及“其他”（包括所有其余死因）。表2按ICD-10分类系统汇总了死因分组的信息。

表2：ICD编码及六大类别分类。

标题	**ICD 10**	**死因**
传染性疾病	A00–B99	传染性疾病
肿瘤	C00–D48	肿瘤
心脏病	I00–I52	CVD
脑血管疾病	G45, I60–I69	CVD
循环系统其他及未特指疾病	I70–I99	CVD
急性呼吸系统疾病	J00–J22, U04	呼吸系统
其他呼吸系统疾病	J30–J98	呼吸系统
外因	V01–Y98	外部
所有其他原因	其余 ICD 编码	其他

采用1989–2008年的数据（法国采用1989–2007年的数据）来估计模型参数的后验分布，其余数据（美国为2009–2023年，法国为2008–2022年）则作为留出集，用于评估所提方法的预测性能；两个国家均采用15年窗口。在第二项建模分析中（其结果报告于附录A），还使用1979年至1988年的数据拟合模型，并再次评估美国2009–2023年和法国2008–2022年这一时期的预测性能。

第 3.3节所述的 HMC采样器并行运行了三次，生成了三条链，每条链包含 3,000 个样本，其中前 1,500 次抽样用作预热期。我们通过表 $\hat{}$中的 R 诊断结果评估了15和16所示的参数后验分布的收敛情况（例如，参见 Gelman和Rubin，1992年；Stan开发团队，2025年）。

6 结果

^首先，对于大多数参数， R 低于建议的 1.05 阈值，这表明三条链已收敛至平稳后验分布。

附录中列出了对所分析的两个国家按性别和样本内年份划分的贝叶斯 AP 和 LC 模型标量参数的后验汇总 C。

令 $\log \widehat{m}_{a,t,c}$ 表示 $\log m_{a,t,c}$ 的预测值，即所考虑模型下年龄 a、年份 t、死因 c 的死亡率对数值。通过计算评估集 $\mathcal{J}$（样本外年份中所有 (a,t,c)）的平均绝对误差（MAE）和均方根误差（RMSE），评估其准确性：

$$\mathrm{MAE} = \frac{1}{|\mathcal{J}|} \sum_{(a,t,c)\in\mathcal{J}} \left|\log m_{a,t,c} - \log \widehat{m}_{a,t,c}\right|,$$

$$\mathrm{RMSE} = \left[\frac{1}{|\mathcal{J}|} \sum_{(a,t,c)\in\mathcal{J}} \left(\log m_{a,t,c} - \log \widehat{m}_{a,t,c}\right)^2\right]^{1/2}.$$

这两项指标也按死因分别计算，并与相应的样本内汇总值一并列出；随后将其与按单个死因拟合的 LC 模型以及 CoDA 方法所得结果进行比较。这两种竞争方法的主要特点概述于附录 B。

6.1 美国（美国）

在滚动 15 年样本外评估中，对 DMP-AP、DMP-LC、CoDa 和 LC 进行比较。表3按性别列出了死亡总数的样本内（训练）和样本外（OOS）平均绝对误差及均方根误差。

总死亡率表3 显示，在总死亡率曲面上，DMP-LC 设定具有明显优势。对于女性和男性，DMP-LC 的样本内和样本外误差均最低。这与潜在时期指数的作用相符：它能够捕捉主要的长期趋势，同时借助平滑的年龄结构在不同年龄之间共享信息。在相干贝叶斯设定中，DMP-AP 通常排名第二，表明加性年龄–

时期结构仍具有竞争力。LC 分解在本应用中对时间趋势的刻画更为鲜明。

表 3：美国——按性别划分的总死亡率表现（训练 vs 样本外）。越低越好。最佳以**粗体**标示，次佳加下划线（各性别内）。

Sex	模型	训练均方根误差	训练平均绝对误差	样本外均方根误差	样本外平均绝对误差
女性	DMP-AP	0.0084	0.0027	0.0208	0.0072
	DMP-LC	**0.0047**	**0.0015**	**0.0090**	**0.0039**
	CoDa	0.1475	0.0833	0.1464	0.0808
	LC	0.0101	0.0034	0.0134	0.0047
男性	DMP-AP	0.0174	0.0058	0.0199	0.0074
	DMP-LC	**0.0038**	**0.0015**	**0.0110**	**0.0051**
	CoDa	0.1690	0.1011	0.1663	0.0944
	LC	0.0158	0.0055	0.0179	0.0070

相比之下，LC基准（在我们的实现中对每个序列分别拟合）在总量上的准确度始终低于DMP-LC，尽管在若干单元格的总体指标上，它仍是表现第二好的方法。这表明，采用显式平滑先验和联合推断的贝叶斯建模框架，相较于经典LC基线方法确实有所提升。在这项评估中，CoDa 方法在总死亡率上的误差明显更大；这并不意外，因为CoDa方法旨在生命表（单纯形）几何结构中对死亡分布建模，而非直接对死亡率建模。

按死因划分的死亡率 表 4 展示了按性别和死因划分的样本外均方根误差和平均绝对误差。不同死因的结果各异，但有两种经验模式值得注意。DMP-AP和DMP-LC模型在若干主要死因上始终名列前茅（排名第一或第二），同时保持总量与各死因之间的内部相干性。具体而言，DMP-LC在两种性别的肿瘤（NEOP）方面都具有竞争力，并且在男性心血管疾病（CVD）的均方根误差指标上表现突出。与此同时，从平均绝对误差来看，DMP-AP在心血管疾病（CVD）方面仍具竞争力；在女性外因（EXT）方面也表现不俗。这些发现与模型设计初衷一致。事实上，DMP-LC使用单一潜在时间因子，同时驱动整体死亡水平以及各死因之间的系统性重新分配。相比之下，DMP-AP允许总量和构成都采用更灵活的（加性）时期成分。

此外，c 分死因潜类别和 CoDa 方法在分析特定

死因时可能表现更好，但这些优势并不会一致地延伸至整个系统。这一区别很重要：具体而言，若按死因独立拟合潜类别模型（这是一种常见做法），并不能保证分死因预测之和与总死亡率预测之间具有相干性，因此可能产生内部不一致的预测。Kjærgaard 等人（2019年）讨论了独立预测死因别死亡率这一做法的普遍局限，并提出了尊重竞争风险结构的相干替代方案。类似地，CoDa 方法通过构造在组成数据（生命表死亡数）空间中确保相干性，但它们处理的是经过变换的生命表量（例如经过中心对数比（CLR）变换的生命表死亡数），并且需要额外的分析步骤；这些特点可能使其表现容易受到具体评估目标（死亡率还是生命表死亡数）以及数据特性的影响，尤其是对于较小的死因。

表4：美国——按死因和模型划分的样本外表现，性别作为列分组显示。RMSE/MAE 越低越好。在每个死因 × 性别内，分别针对 RMSE 和 MAE，最佳结果以**粗体**显示，次佳结果加下划线。

		女性		男性	
死因	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
CVD	DMP-AP	**0.0060**	0.0024	0.0132	**0.0042**
	DMP-LC	0.0134	0.0056	**0.0112**	0.0052
	CoDa	0.0066	**0.0020**	0.0146	0.0050
	LC	0.0158	0.0068	0.0124	0.0061
EXT	DMP-AP	**0.0004**	**0.0002**	0.0008	0.0004
	DMP-LC	0.0004	0.0002	**0.0004**	0.0003
	CoDa	0.0004	0.0002	0.0007	0.0004
	LC	0.0017	0.0007	0.0005	**0.0003**
INF	DMP-AP	0.0006	0.0002	0.0012	0.0006
	DMP-LC	0.0005	0.0002	0.0008	0.0004
	CoDa	**0.0002**	**0.0001**	0.0012	0.0005
	LC	0.0004	0.0002	**0.0004**	**0.0002**
NEOP	DMP-AP	0.0024	0.0011	0.0030	0.0013
	DMP-LC	**0.0007**	**0.0004**	0.0023	0.0013
	CoDa	0.0011	0.0005	**0.0017**	**0.0007**

续见下页

		女性		男性	
死因	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
	LC	0.0008	0.0005	0.0023	0.0014
其他	DMP-AP	0.0081	0.0032	0.0080	**0.0029**
	DMP-LC	0.0100	0.0035	**0.0076**	0.0030
	CoDa	**0.0069**	**0.0023**	0.0095	0.0036
	LC	0.0279	0.0105	0.0310	0.0127
RESP	DMP-AP	0.0080	0.0026	0.0058	0.0021
	DMP-LC	0.0044	0.0016	0.0056	0.0022
	CoDa	0.0035	**0.0010**	**0.0044**	**0.0017**
	LC	**0.0033**	0.0012	0.0046	0.0019

表 5 汇总了按死因计算的RMSE和平均绝对误差的算术平均值。男性的RMSE以DMP-LC模型最低，而平均绝对误差以DMP-AP模型最低。这些结果表明，在汇总各死因的两项预测性能指标时，连贯的贝叶斯方法总体上实现了最佳权衡。女性方面，CoDa模型在各死因上的平均误差最低。不过，这一结果应谨慎看待。在总体分析中，采用相同方法时，按当前基于率的指标衡量，全因死亡率的表现较差。这表明，组成生命表模型优化的目标与此处评估的内容并不相同。

表5：美国—按性别划分的各死因样本外平均表现。数值越低越好。最佳结果以**粗体**显示，次佳结果加下划线（各性别内比较）。

Sex	**模型**	**均方根误差均值**	**平均绝对误差均值**
女性	DMP-AP	0.0042	0.0016
	DMP-LC	0.0049	0.0019
	CoDa	**0.0031**	**0.0010**
	LC	0.0083	0.0033
Male	DMP-AP	0.0053	**0.0019**
	DMP-LC	**0.0047**	0.0021
	CoDa	0.0054	0.0020
	LC	0.0085	0.0038

相比之下，DMP -潜类别能够给出最准确的总死亡率预测，同时仍保持

在多种死因的预测中具有竞争力，这正是当死亡率水平及其构成分解都与应用相关时，一致性预测的目标。

这些结果支持两项结论。1）在预测期内，贝叶斯 LC–DM 形式对两种性别的总死亡率预测最为准确，同时从构造上保证了总死亡率与分死因死亡率之间的相干性。2）尽管某些基准方法在特定死因的预测上优于我们的模型，但它们无法同时实现：（i）准确预测总量；（ii）相干分解；以及（iii）通过单一联合概率模型，对水平和构成都给出预测不确定性。这种联合且相干的视角对公共卫生规划至关重要，因为死因预测在内部不一致时，即使各死因的变化轨迹看似合理，也可能导致不合乎情理的分解（Kjærgaard 等人，2019年）。

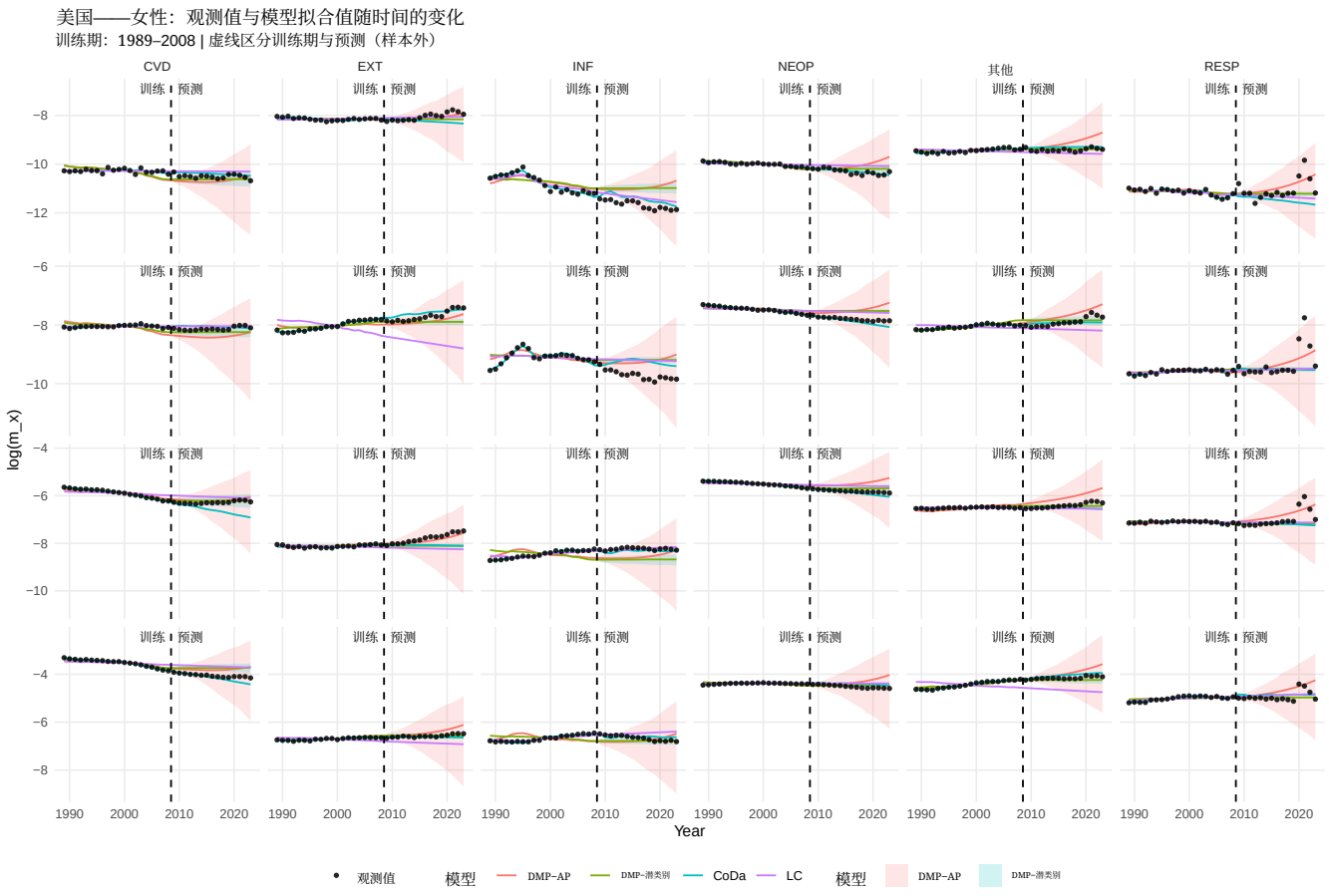

图 1：美国女性：选定年龄（20、40、60、80）的分死因观测值（黑点）以及拟合/预测对数死亡率。竖直虚线标示样本内时期（1989–2008）的结束和样本外预测的开始。

图1和2 展示了选定年龄（20、

40、60、80 岁）各死因对数死亡率随时间的观测值，竖直虚线标示样本内窗口（1989年–2008年）向预测窗口的过渡。观测死亡率（黑点）通常与两种相干的 DMP 规格（DMP-AP 和 DMP-LC）拟合良好。2008年之后，其预测区间逐渐变宽，反映出预测期内不确定性不断增加。这种现象在变异性较高和/或年轻年龄段信号较弱的死因中尤为明显，在这些情况下，样本外的可信区间大幅扩展。

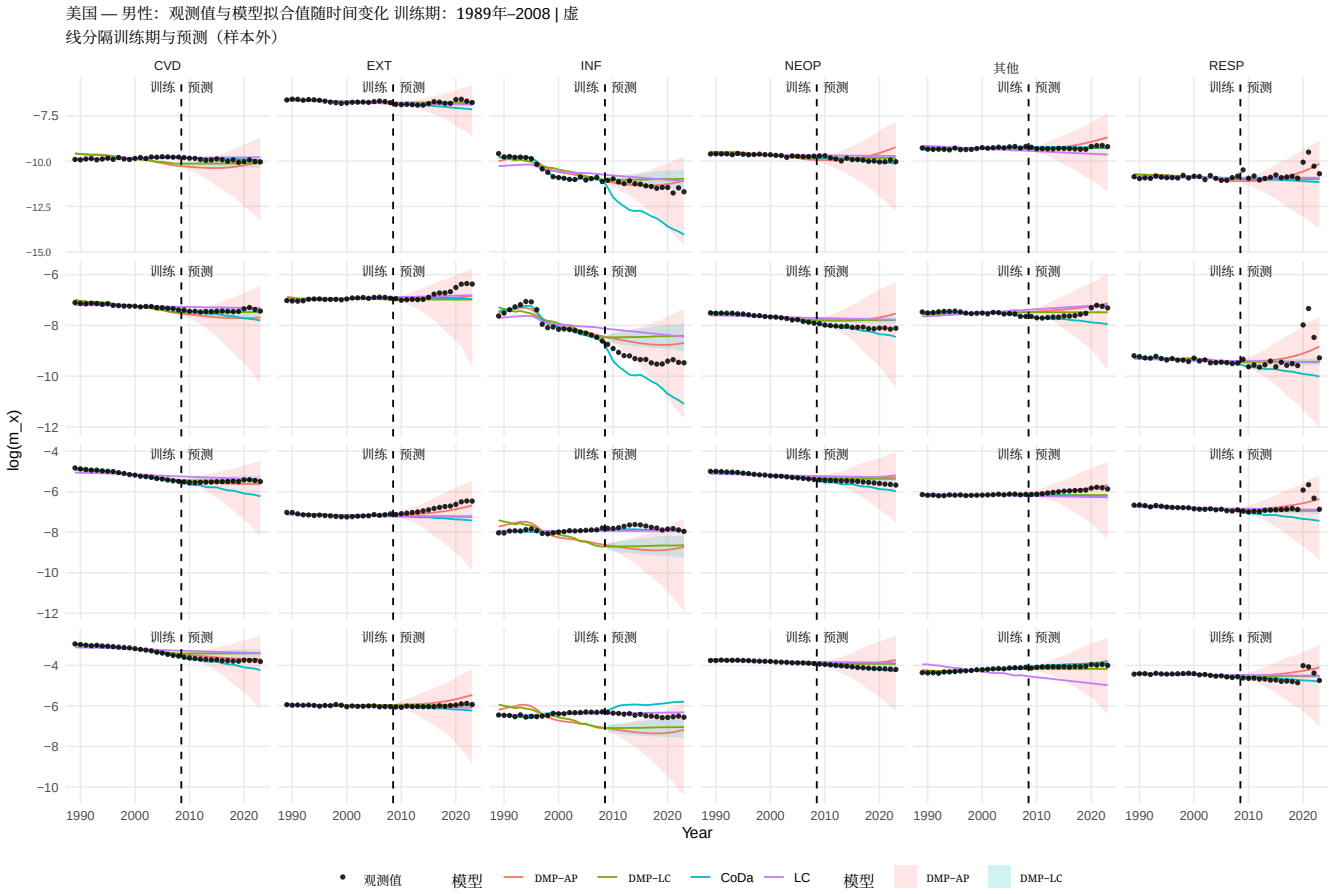

图2：美国男性：选定年龄（20、40、60、80岁）各死因的观测值（黑点）及拟合／预测对数死亡率。竖直虚线标示样本内时期（1989–2008）的结束和样本外预测的开始。

在两种性别中，DMP-LC 往往能在以平滑趋势为主的面板（例如较高年龄组的心血管疾病和肿瘤）中提供最稳定的外推；而当 2008 年后的动态变化与 2008 年前加性时期模式的延续相似时，DMP-AP 仍具竞争力。相比之下，非相干基准模型在样本外可能呈现更显著的面板特异性偏离。单死因LC拟合有时会在某些年龄—死因组合中相对于观测值出现系统性漂移；CoDa 则在某些传染性疾病／呼吸系统面板中偶尔出现较大偏差（尤其是在较低年龄组），这与其在基于率轨迹进行评估时对生命表／组成数据表示的敏感性一致，

在死亡率轨迹上进行评估时的表现。

总体而言，图表支持此前的发现：相干贝叶斯模型在样本内校准方面表现出色，样本外轨迹合理且更具竞争力，并且不确定性区间在样本内边界之外适当地扩展，同时在每个年龄和时期都保持内部一致的死因分解。

6.2 法国

我们采用与美国相同的评估设计报告法国的结果：将最后15个日历年作为样本外数据留出，此前的年份则构成样本内数据集。我们使用均方根误差和平均绝对误差（数值越低越好），分别按性别报告（见表 6–8）。

表 6 显示，当样本内时期始于1989年时，DMP-LC规格在法国两种性别的总死亡率预测中均占主导地位。女性方面，DMP-LC的样本内和样本外误差均为最低，且显著优于其他方法。一个显著特点是，DMP-AP对女性总量的样本外误差非常大，这表明在预测期内对水平分量的外推不稳定。男性方面，DMP-LC在总量预测上仍然是最准确的方法，其次是经典的潜类别基准；而在这项总体评估中，CoDa在率尺度上的竞争力仍明显较弱。

表6：法国——按性别划分的总死亡率表现（训练与样本外），训练期始于1989年。数值越低越好。每项指标分别按性别评选最佳（**粗体**）和次佳（下划线）。

Sex	模型	训练均方根误差	训练平均绝对误差	样本外均方根误差	样本外平均绝对误差
女性	DMP-AP	0.0092	0.0027	0.2826	0.0647
	DMP-LC	**0.0036**	**0.0010**	**0.0074**	**0.0028**
	CoDa	0.1527	0.0838	0.1413	0.0738
	LC	0.0057	0.0022	0.0094	0.0046
男性	DMP-AP	0.0160	0.0047	0.0324	0.0121
	DMP-LC	**0.0078**	**0.0022**	**0.0100**	**0.0045**
	CoDa	0.1784	0.1048	0.1624	0.0910
	LC	0.0087	0.0034	0.0129	0.0069

表7 按死因详细列出了样本外表现。整体情况各异，但呈现出两种模式。首先，对于两种性别的大多数死因，DMP-LC始终是表现最佳的模型之一，

同时从构造上确保总量与死因分解之间的相干性。具体而言，在男性中，DMP-LC在心血管疾病（CVD）、外因（EXT）和残余类别（OTHER）的均方根误差（RMSE）和平均绝对误差（MAE）方面表现最佳，在肿瘤（NEOP）方面也极具竞争力。在女性中，DMP-LC在EXT和OTHER方面表现最佳，在CVD和传染病（INF）方面也接近最佳，这证实单一潜在时间指标能够同时捕捉主导趋势以及各死因之间的系统性重新分配。

表7：法国——按死因和模型划分的样本外表现，自1989年起训练，性别显示为列组。均方根误差/平均绝对误差越低越好。最佳以**粗体**标示，次佳加下划线（在每个死因 ×性别内，分别针对均方根误差和平均绝对误差）。

		女性		男性	
死因	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
CVD	DMP-AP	0.0752	0.0178	0.0117	0.0046
	DMP-LC	0.0055	0.0023	**0.0064**	**0.0029**
	CoDa	**0.0038**	**0.0010**	0.0129	0.0039
	LC	0.0094	0.0042	0.0092	0.0045
EXT	DMP-AP	0.0217	0.0047	0.0046	0.0015
	DMP-LC	**0.0008**	**0.0003**	**0.0013**	**0.0004**
	CoDa	0.0017	0.0005	0.0016	0.0005
	LC	0.0011	0.0005	0.0013	0.0005
INF	DMP-AP	0.0052	0.0012	0.0016	0.0006
	DMP-LC	**0.0004**	0.0002	0.0013	0.0005
	CoDa	0.0007	0.0002	0.0017	0.0005
	LC	0.0005	**0.0002**	**0.0005**	**0.0002**
NEOP	DMP-AP	0.0462	0.0123	0.0085	0.0029
	DMP-LC	0.0008	0.0003	**0.0020**	0.0011
	CoDa	0.0015	0.0007	0.0028	**0.0010**
	LC	**0.0007**	**0.0003**	0.0024	0.0015
其他	DMP-AP	0.0932	0.0210	0.0086	0.0029
	DMP-LC	**0.0033**	**0.0009**	**0.0022**	**0.0008**
	CoDa	0.0038	0.0012	0.0053	0.0016
	LC	0.0047	0.0014	0.0031	0.0011

续见下页

		女性		男性	
死因	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
RESP	DMP-AP	0.0443	0.0080	**0.0046**	**0.0014**
	DMP-LC	**0.0043**	**0.0011**	0.0071	0.0019
	CoDa	0.0058	0.0015	0.0120	0.0033
	LC	0.0043	0.0012	0.0069	0.0021

其次，基准方法在某些特定死因上的表现可能很强，但并非始终占优。例如，CoDa 在女性心血管疾病（RMSE 和 MAE）方面表现最佳，在其他几种女性死因上的表现也颇具竞争力；而经典 LC 基准在女性肿瘤和 INF （尤其是 MAE ）方面表现最佳，在男性 INF 方面也颇具竞争力。与美国数据的分析一样，解读这类逐死因的优势时，还应结合系统层面的一致性：分别拟合的 LC 模型（每个死因一个模型）并不能保证分死因预测之和等于总死亡率的预测，而两种贝叶斯设定都能确保各年龄和时期上的相干性。

表 8 中各死因的平均值进一步印证了这些结论。男性方面，DMP-LC 的跨死因平均 RMSE 和平均 MAE 最低。女性方面，DMP-LC 的平均 RMSE 最低，CoDa 的平均 MAE 最低；而 DMP-AP 则是一个异常值，因为它在多个死因上，尤其是在总量方面，误差都非常大。

表8：法国——各死因的样本外均值表现，自1989年起训练，按性别划分。数值越低越好。最佳以**粗体**标示，次佳加下划线（在各性别内，分别针对均方根误差均值和平均绝对误差均值）。

Sex	**模型**	**均方根误差均值**	**平均绝对误差均值**
女性	DMP-AP	0.0476	0.0108
	DMP-LC	**0.0025**	0.0009
	CoDa	0.0029	**0.0008**
	LC	0.0034	0.0013
Male	DMP-AP	0.0066	0.0023
	DMP-LC	**0.0034**	**0.0013**
	CoDa	0.0061	0.0018
	LC	0.0039	0.0016

图 3 和 4 通过展示选定年龄（20、40、60、80）的观测对数死亡率、拟合轨迹和预测行为，对这些结果进行了直观检验，虚线

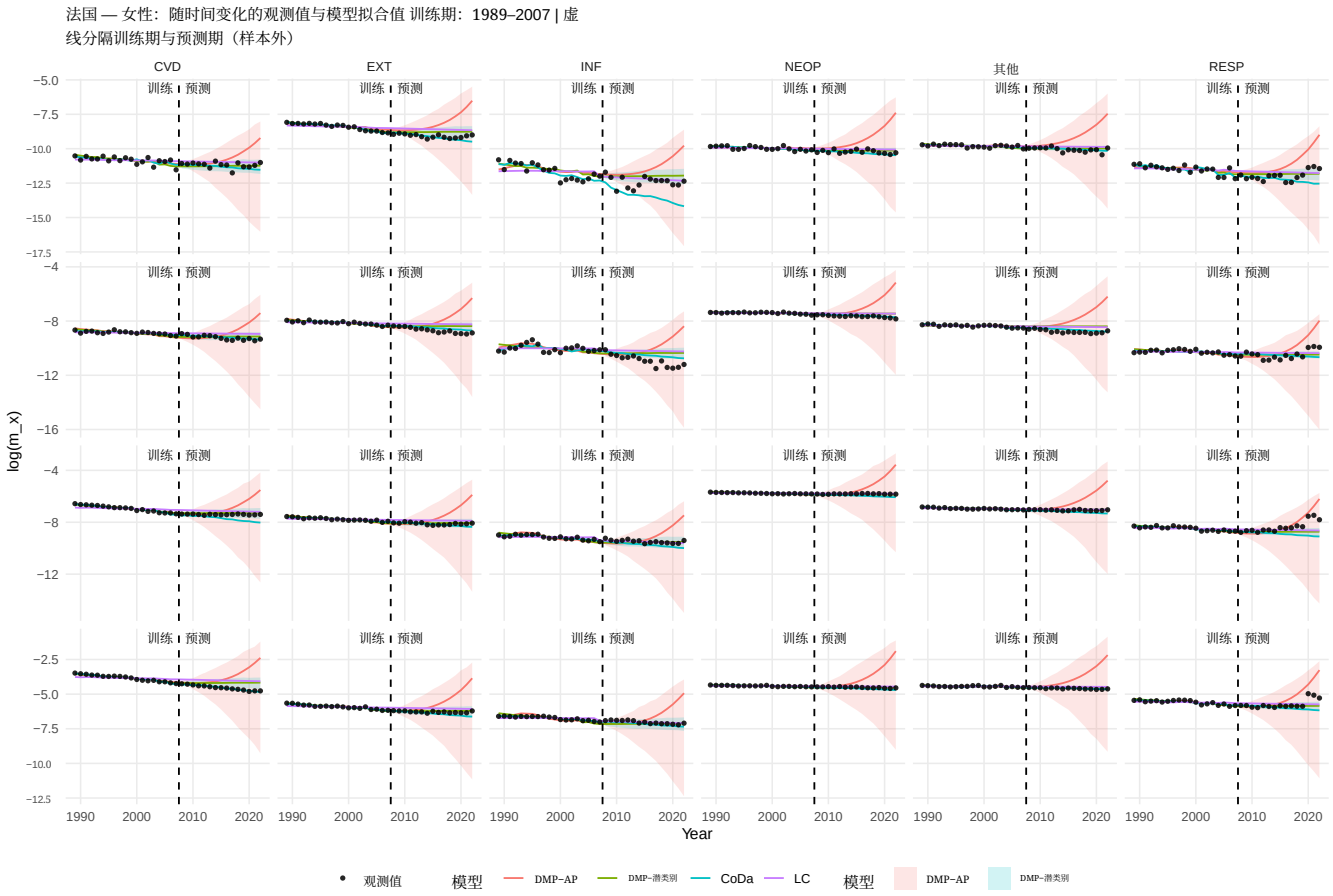

图3：法国女性：选定年龄（20、40、60、80）各死因的对数死亡率观测值（黑点）、拟合值和预测值。竖直虚线标示样本内时期（1989–2007）的结束及样本外预测的开始。

表示从样本内（1989–2007）到样本外预测（2008年起）的过渡。图中显示，DMP-LC 紧密跟踪 2008 年以前的趋势，并在 2007 年之后产生稳定的外推结果；随着预测期延长，不确定性区间逐渐变宽。相比之下，DMP-AP 呈现出非常宽的预测区间，并在多个面板中出现明显漂移（女性尤为突出），这与其在总量及多个死因上的样本外表现较差相一致。这种行为符合时期成分高度灵活的特征，其二阶随机游走外推可能在较长预测期内产生迅速增加的不确定性和曲率。CoDa 和潜类别模型在部分死因–年龄面板（尤其是 INF 和肿瘤）中仍具竞争力，但它们的优势并未均匀覆盖完整的死因划分，也无法像相干贝叶斯建模框架那样，同时为水平和构成提供准确的总量、相干分解以及联合概率不确定性的组合。

附录A中的表9–14 展示了采用更长的样本内窗口进行的评估，该窗口始于

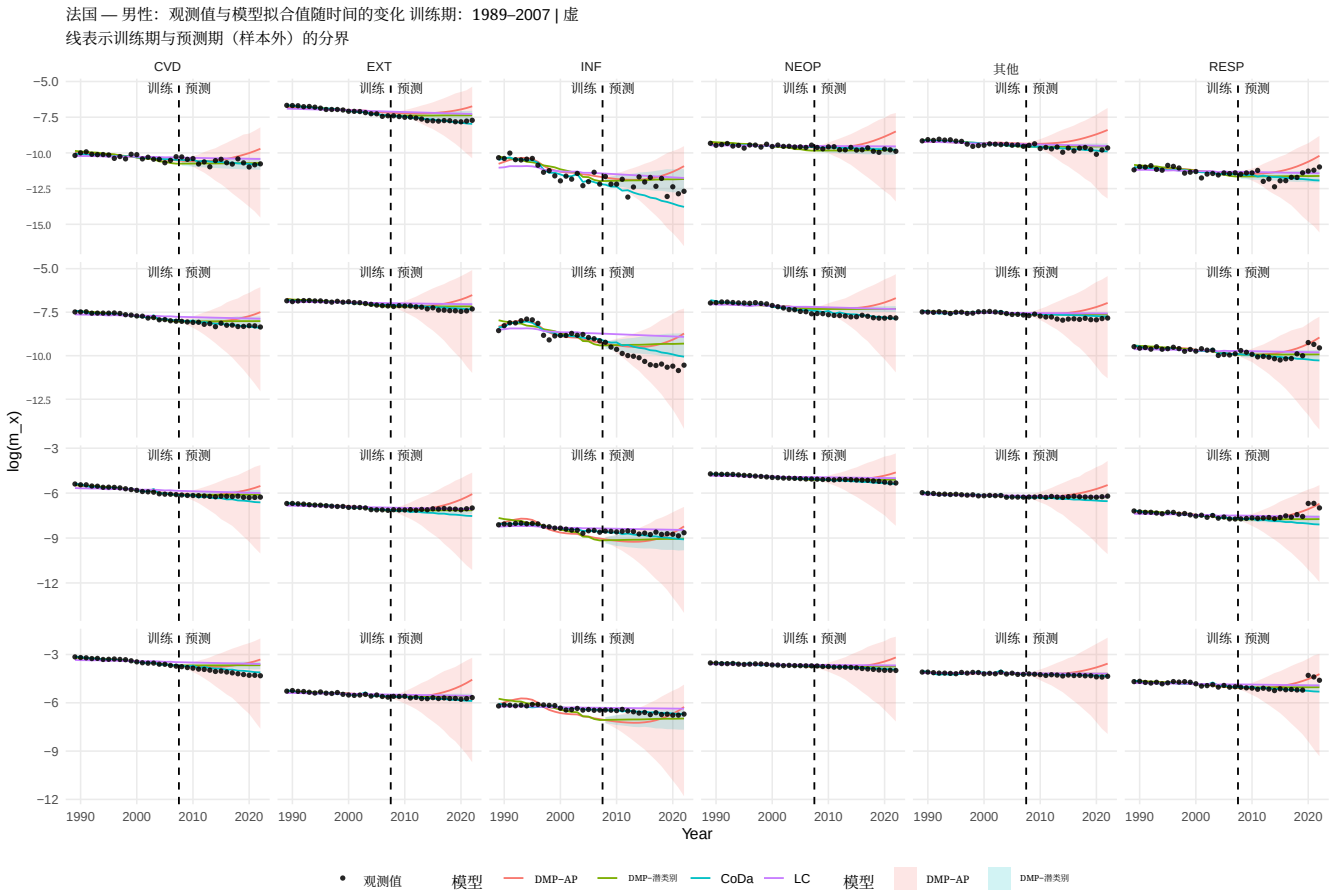

图 4：法国男性：选定年龄（20、40、60、80）的观测值（黑点）以及按死因划分的拟合/预测对数死亡率。竖直虚线标示样本内时期（1989–2007）的结束和样本外预测的开始。

1979年。对于两个国家，DMP-LC 在两种性别的总死亡率预测方面仍表现最佳，这证实了扩大估计样本后，基于 LC 的相干贝叶斯设定依然稳健。相比之下，尽管 CoDa 偶尔在特定死因上表现出色，但在这项基于率的评估中，其整体竞争力仍明显较弱。

就死因特异性表现而言，较长的样本内窗口放大了正文中已观测到的异质性。在美国数据中，CoDa 在各死因上的平均误差通常最低（表 13），而 DMP-AP/DMP-LC 的表现始终具有竞争力，且在总量预测上通常占优。法国也呈现出相同的模式。具体而言，女性的各死因平均误差以 CoDa 最低，而 DMP-LC 对男性的总体表现最佳，并且在按死因评估时也名列前茅（表 14）。总体而言，附录结果进一步印证了主要结论：相干贝叶斯模型，尤其是 DMP-LC，在保持构成结构相干的同时，能为总死亡率提供最可靠的准确性；而基准方法可能在

某些特定死因上表现出色，但其系统层面的表现并不始终如一。

总体而言，从1979年开始的估计证实了两项稳健发现：1）DMP-LC在不同国家和性别中的总死亡率预测最稳定、最准确，而且即使大幅延长估计窗口，其优势依然存在。2）各死因的预测表现存在异质性，并取决于评估目标。具体而言，CoDa在若干单个死因上表现出色，在某些情境下（如女性人口）甚至能使各死因的误差最小；但其总量预测表现不佳。这进一步支持了所提出的方法，强调相干贝叶斯联合模型（DMP-AP，尤其是DMP-LC）能够提供更准确的预测总量、相干分解，以及对不确定性的统一概率描述。相比之下，基准方法在某些死因上的收益未必能推广到整个竞争风险／组成系统。

7 结论

本文提出了一个用于分析和预测按死因划分的死亡率的狄利克雷–多项式–泊松建模框架。所提出的层级设定提供了一种模块化结构，将总死亡率随年龄和日历年变化的情况，与死亡在各死因间的分配联合建模。总死亡人数和死因特异性概率的随机性，使模型能够捕捉总体死亡率以及不同死因相对重要性的年龄和时间依赖动态。

通过同时对这些组成部分进行建模，该框架能够提供相干的死因特异性死亡率预测，并确保其与全因死亡率预测完全一致。重要的是，这种相干性不仅体现在点预测层面，也体现在相关的不确定性上，因为模型明确表示了死因间的依赖关系。由此得到的方法为独立的死因特异性模型提供了一种易于处理且便于解释的替代方案，同时保留了分析透明度，并与贝叶斯推断相兼容。

这些优势也伴随着代价。对年龄、时期和死因各维度进行完整的贝叶斯估计，计算成本可能很高，并且需要谨慎设定先验并进行诊断评估；此外，长期预测可能对所假定的时间结构较为敏感。从实证角度看，针对美国和法国的回测分析凸显了以下方面的重要性：

相干联合建模。基于 LC 的相干贝叶斯设定（DMP–LC）在不同国家和性别的总体死亡率曲面上，取得了最低或接近最低的样本外误差，同时在较长期限内保持稳定。尤其是在法国，潜在时期指数在捕捉主导性长期趋势以及在相干多原因系统内跨年龄借力方面发挥着关键作用。在原因特定层面，DMP–LC 在大多数死因和人群中均取得较低的均方根误差，避免了系统性偏差，并保留了连接各死因与总量的聚合约束。其他方法可能在个别情况下表现更佳，但无法同时保证总体与分项预测之间的相干性，也无法提供总量和各死因的联合预测分布。当内部一致性和可解释的预测不确定性至关重要时，相干贝叶斯框架在准确性、相干性和不确定性量化之间实现了良好的权衡。

有几项扩展值得进一步研究。首先，可以允许狄利克雷浓度参数随年龄或时期变化，从而使组成部分的离散度具有更大的灵活性。其次，可以引入额外的过度离散或队列结构，以评估这些特征是否能在时期动态之外进一步提升预测价值。虽然这两项扩展均可行，但是否采用应由谨慎的成本效益评估来决定：只有当额外的计算和建模复杂性能够在预期应用中切实改善校准、相干性或预测准确度时，才值得引入。

致谢

Andrea Nigri感谢澳大利亚国立大学人口学院以及麦考瑞大学精算研究与商业分析系研讨会参与者提出的富有洞见的意见和建议。

参考文献

U. Basellini、C. G. Camarda和H. Booth。三十年回顾：用于死亡率预测的Lee–Carter方法综述。《国际预测杂志》，39(3):1033–1049，2023年。H. Booth和L. Tickle。死亡率建模与预测：方法综述。《精算科学年刊》，3(1-2):3–43，2008年。N. Brouhns、M. Denuit和J. K. Vermunt。构建预测生命表的泊松对数双线性回归方法。《保险：数学与经济学》，31(3):373–393，2002年。J. Bryant和J. L. Zhang。《贝叶斯人口估计与预测》。Chapman & Hall/CRC出版社，博卡拉顿，佛罗里达州，2018年。G. Cardillo、P. Giordani、S. Levantesi、A. Nigri和A. Spelta。使用四路CANDECOMP/PARAFAC分解进行死亡率预测。《斯堪的纳维亚精算杂志》，2023(9): 916–932，2023年。G. Cardillo、P. Giordani、S. Levantesi和A. Nigri。基于张量的死因死亡率建模方法。《运筹学年刊》，342(3):2075–2094，2024年。C. Chatfield。《时间序列预测》。Chapman & Hall，博卡拉顿，2000年。C. L. Chiang。死亡率分析中的竞争风险。《公共卫生年度评论》，12(1):281–307，1991年。

E. M. Crimmins. 寿命与健康寿命：过去、现在与未来。《老年学家》，55(6): 901–911，2015年。

A. Gelman 和 D. B. Rubin. 使用多条序列进行迭代模拟推断。《统计科学》，7(4):457–472，1992年。

R. Graziani 和 A. Nigri. 狄利克雷框架下的年龄–时期–队列模型：一致的死因估计。《保险：数学与经济学》，126:103194，2026年。

M. D. Hoffman 和 A. Gelman。《无转向采样器：在哈密顿蒙特卡洛中自适应设定路径长度》。《机器学习研究杂志》，15(1):1593–1623，2014。

人类死因数据系列。人类死亡率数据库。马克斯·普朗克人口研究所（德国）、加利福尼亚大学伯克利分校（美国）和法国人口研究所（法国），2026年。可在www.mortality.org 获取（数据下载于2024年6月26日）。

人类死亡率数据库。 加利福尼亚大学伯克利分校，美国；马克斯·普朗克人口研究所，德国，2026年。可在以下网址在线获取 http://www.mortality.org[访问日期：6月26日-2024]。

A. Hunt 与 D. Blake。构建死亡率模型的一般程序。《北美精算学杂志》，18(1):116–138，2014。

R. J. Hyndman、H. Booth 与 F. Yasmeen。一致性死亡率预测：基于函数型时间序列模型的乘积比方法。《人口学》，50:261–283，2013。

S. Kjærgaard、Y. E. Ergemen、M. Kallestrup-Lamb、J. Oeppen 与 R. Lindahl-Jacobsen。利用组成数据分析预测死因：以癌症死亡为例。《英国皇家统计学会杂志*C*辑：应用统计》，68(5):1351–1370，2019年。

G. Lanfiuti Baldi、A. Nigri 和 H. L. Shang。死亡率差异的年龄—时期建模：癌症和循环系统疾病案例。《官方统计学杂志》，41(4):1154–1180，2025年。

R. D. Lee 和 L.R. Carter。美国死亡率的建模与预测。《美国统计学会杂志：应用与案例研究》，87(419):659–671，1992年。

H. Li、H. Li、Y. Lu 和 A. Panagiotelis。死因死亡率建模的预测协调方法。《保险：数学与经济学》，86:122–133，2019年。

J. E. Mosimann。论复合多项分布、多变量 β 分布以及比例间相关性。《生物计量学》，49(1 and 2):65–82，1962年。

J. Oeppen。多重减员生命表的协调预测：使用日本死因数据进行的检验。载于 欧洲人口研究协会，2008年。网址：https://epc2008.eaps.nl/papers/80611。

S. Preston、P. Heuveline 和 M. Guillot。《人口学：人口过程的测量与建模》。布莱克威尔出版社，马尔登，马萨诸塞州，2000年。

A. E. Renshaw 和 S. Haberman。针对死亡率降低因子的李–卡特模型的队列扩展。《保险：数学与经济学》，38(3):556–570，2006年。

Stan开发团队。Stan建模语言用户指南和参考手册。
https://mc-stan.org/docs/stan-users-guide/index.html，2025年。版本 2.38。

J. R. Wilmoth。按死因分解后，死亡率预测是否总是更加悲观？《数学人口研究》，5(4):293–319，1995年。

T. 郑和 R. 陈。用于组成数据时间序列的狄利克雷 ARMA 模型。《多元分析杂志》，158：31–46，2017年。

美国和法国，1979年

表9–13 和10–14 报告了一项稳健性检验，其中样本内窗口向前延伸至1979年（保持相同的15年样本外预测期）。这项实验很有参考意义，因为较长的估计样本涵盖了更多流行病学阶段和结构性变化，从而对各竞争模型的外推表现进行了压力测试。

表9：美国——总死亡率表现（训练与样本外），按性别划分，训练数据始于1979年。数值越低越好。最佳结果用**粗体**表示，次佳结果加下划线（在每种性别内，按各项指标分别比较）。

		训练		OOS	
Sex	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
女性	DMP-AP	0.0132	0.0039	0.0184	0.0069
	DMP-LC	**0.0044**	**0.0014**	**0.0101**	**0.0042**
	CoDa	0.1446	0.0831	0.1458	0.0804
	LC	0.0070	0.0030	0.0406	0.0178
男性	DMP-AP	0.0254	0.0084	0.0290	0.0102
	DMP-LC	**0.0046**	**0.0017**	**0.0127**	**0.0058**
	CoDa	0.1657	0.1016	0.1668	0.0951
	LC	0.1177	0.0257	0.0565	0.0219

表10：法国——按性别划分的总死亡率表现（训练与样本外），训练期始于1979年。数值越低越好。最佳结果以**粗体**显示，次佳结果加下划线（分别在每种性别的各项指标中比较）。

		训练		OOS	
Sex	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
女性	DMP-AP	0.0175	0.0048	0.1376	0.0354
	DMP-LC	**0.0040**	**0.0012**	**0.0074**	**0.0028**
	CoDa	0.1594	0.0894	0.1403	0.0730
	LC	0.0127	0.0056	0.1011	0.0259
男性	DMP-AP	0.0215	0.0068	0.0247	0.0100
	DMP-LC	**0.0082**	**0.0023**	**0.0108**	**0.0047**
	CoDa	0.1869	0.1114	0.1603	0.0889
	LC	0.0161	0.0072	0.0270	0.0157

在两个国家和两种性别中，按总量计算的排名都非常稳定。事实上，DMP-LC 在样本内和样本外始终是整体表现最佳的方法。对于美国（表 9），DMP-LC 对女性和男性的样本外误差均最低，DMP-AP 通常排名第二，而基准方法（潜类别，尤其是 CoDa）在总量预测方面的竞争力则明显较弱

。法国男性的排名也相同（表10）。对于法国女性，DMP-LC 再次占据优势，而 DMP-AP 和潜类别的样本外误差则大得多，这表明即使在使用更长历史窗口进行训练时，加性年龄—时期设定在这种情况下仍可能较为敏感。总体而言，延长样本内时期进一步支持这样的判断：基于 LC 的相干贝叶斯设定能够提供最可靠的系统层面预测，即同时对水平进行建模，并保持相干分解。

在死因层面，较长的样本内窗口凸显了正文中已经显现的异质性。在美国（表 11）中，CoDa 和潜类别模型（LC）在特定死因上可能表现最佳（例如，INF 通常更适合 CoDa/LC；对于女性，EXT 可能更适合 LC），而 DMP-LC 往往具有竞争力，并且常在 CVD 等男性主要死因上表现最佳。在若干单元格中（尤其是女性 CVD），DMP-AP 仍接近最佳方法，这表明加性年龄—时期结构仍能捕捉构成中的重要变化；但若目标是准确预测总量并保持相干性，其系统层面优势并不像 DMP-LC 那样稳健。

表11：美国——按死因和模型划分的样本外表现，训练数据始于1979年，性别以列组显示。RMSE/平均绝对误差越低越好。最佳结果以**粗体**显示，次佳结果加下划线（在每个死因 × 性别组内，分别按RMSE和平均绝对误差比较）。

		女性		男性	
死因	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
CVD	DMP-AP	**0.0049**	0.0021	0.0177	0.0057
	DMP-LC	0.0142	0.0059	**0.0109**	0.0050
	CoDa	0.0063	**0.0019**	0.0145	**0.0041**
	LC	0.0263	0.0126	0.0233	0.0127
EXT	DMP-AP	0.0005	0.0002	0.0007	0.0004
	DMP-LC	0.0004	0.0002	**0.0005**	**0.0003**
	CoDa	0.0007	0.0003	0.0008	0.0004
	LC	**0.0003**	**0.0002**	0.0007	0.0004
INF	DMP-AP	0.0010	0.0003	0.0010	0.0005
	DMP-LC	0.0018	0.0007	0.0018	0.0008
	CoDa	**0.0002**	**0.0001**	0.0007	0.0002
	LC	0.0004	0.0001	**0.0004**	**0.0002**

续见下页

		女性		男性	
死因	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
NEOP	DMP-AP	0.0026	0.0012	0.0027	0.0013
	DMP-LC	0.0012	0.0006	0.0028	0.0015
	CoDa	**0.0006**	**0.0003**	**0.0022**	**0.0010**
	LC	0.0009	0.0006	0.0030	0.0019
其他	DMP-AP	0.0094	0.0037	0.0109	0.0038
	DMP-LC	0.0131	0.0047	0.0091	0.0037
	CoDa	**0.0060**	**0.0021**	**0.0086**	**0.0031**
	LC	0.0126	0.0035	0.0613	0.0251
RESP	DMP-AP	0.0073	0.0024	0.0048	0.0018
	DMP-LC	0.0059	0.0021	0.0062	0.0025
	CoDa	**0.0031**	**0.0010**	**0.0043**	**0.0016**
	LC	0.0048	0.0019	0.0058	0.0018

法国在表12中呈现出类似的模式。对于男性，DMP-LC 在各种死因上总体表现强劲，并在若干主要单元格中表现最佳（例如 CVD 和其他死因）。对于女性，CoDa 在某些死因上极具竞争力（尤其是 CVD 和其他死因），这也反映在各死因的平均表现中。然而，这种死因层面的优势并未转化为良好的总死亡率表现（按率评估），这进一步印证了一个关键点：不同方法可以优化问题的不同方面（水平与构成），而死因层面拟合更优，并不必然意味着总量预测具有更优的系统层面一致性和准确性。

表12：法国——按死因和模型划分的样本外表现，自1979年起训练，性别以列组显示。均方根误差/平均绝对误差越低越好。最优以**粗体**标示，次优加下划线（在每个死因 ×性别内，分别针对均方根误差和平均绝对误差）。

		女性		男性	
死因	**模型**	**RMSE**	**MAE**	**RMSE**	**MAE**
CVD	DMP-AP	0.0466	0.0123	0.0122	0.0045
	DMP-LC	0.0058	0.0025	**0.0069**	**0.0030**

续见下页

		女性		男性	
死因	模型	RMSE	MAE	RMSE	MAE
	CoDa	**0.0031**	**0.0010**	0.0154	0.0049
	LC	0.0229	0.0110	0.0227	0.0119
EXT	DMP-AP	0.0105	0.0024	0.0022	0.0008
	DMP-LC	0.0015	0.0006	**0.0014**	**0.0005**
	CoDa	**0.0013**	**0.0005**	0.0021	0.0007
	LC	0.0039	0.0018	0.0021	0.0013
INF	DMP-AP	0.0027	0.0007	0.0014	0.0005
	DMP-LC	**0.0006**	**0.0002**	0.0007	0.0003
	CoDa	0.0015	0.0005	0.0020	0.0006
	LC	0.0982	0.0139	**0.0006**	**0.0003**
NEOP	DMP-AP	0.0209	0.0064	0.0050	0.0020
	DMP-LC	0.0013	0.0006	0.0028	0.0014
	CoDa	0.0010	**0.0006**	**0.0027**	**0.0009**
	LC	**0.0010**	0.0006	0.0038	0.0024
其他	DMP-AP	0.0459	0.0114	0.0064	0.0025
	DMP-LC	0.0058	0.0019	**0.0029**	**0.0010**
	CoDa	**0.0026**	**0.0009**	0.0079	0.0029
	LC	0.0092	0.0033	0.0042	0.0020
RESP	DMP-AP	0.0128	0.0026	**0.0059**	**0.0015**
	DMP-LC	**0.0042**	**0.0012**	0.0073	0.0020
	CoDa	0.0065	0.0017	0.0097	0.0026
	LC	0.0044	0.0017	0.0072	0.0026

表 13 和 14进一步阐明了这些权衡。在美国，CoDa 在女性的平均均方根误差/平均绝对误差方面表现最低，在男性的平均指标方面也表现最佳，而 DMP-AP/DMP-LC 仍是有力的竞争者。在法国，CoDa 在女性各死因的平均误差方面最低，而 DMP-LC 在男性中总体表现最佳。这些结果表明，当人们进行汇总时

在各种死因中，组成数据方法可能表现出色；然而，总量却呈现出不同的结果，而相干贝叶斯潜类别模型是唯一一种在整个系统上始终表现强劲的方法（总量准确且分解相干）。

表 13：美国——各死因的样本外均值表现，训练数据截至1979年，按性别划分。数值越低越好。在每种性别中，分别就均方根误差均值和平均绝对误差均值而言，最佳结果以**粗体**显示，次佳结果加下划线。

Sex	模型	均方根误差均值	平均绝对误差均值
女性	DMP-AP	0.0043	0.0017
	DMP-LC	0.0061	0.0024
	CoDa	**0.0028**	**0.0010**
	LC	0.0076	0.0031
男性	DMP-AP	0.0063	0.0023
	DMP-LC	0.0052	0.0023
	CoDa	**0.0052**	**0.0017**
	LC	0.0158	0.0070

表 14：法国——样本外各死因均值表现，训练数据始于1979年，按性别划分。数值越低越好。最佳结果以**粗体**标示，次佳结果加下划线（分别在各性别组内比较均方根误差均值和平均绝对误差均值）。

Sex	模型	均方根误差均值	平均绝对误差均值
女性	DMP-AP	0.0232	0.0060
	DMP-LC	0.0032	0.0012
	CoDa	**0.0027**	**0.0009**
	LC	0.0233	0.0054
男性	DMP-AP	0.0055	0.0020
	DMP-LC	**0.0037**	**0.0014**
	CoDa	0.0066	0.0021
	LC	0.0067	0.0034

图5-8 通过展示选定年龄（20、40、60、80）的观测对数死亡率、拟合轨迹和预测行为，对这些结果进行直观检验；虚线标示了从样本内时期（1979-2008）到样本外预测（2008年起）的转换。

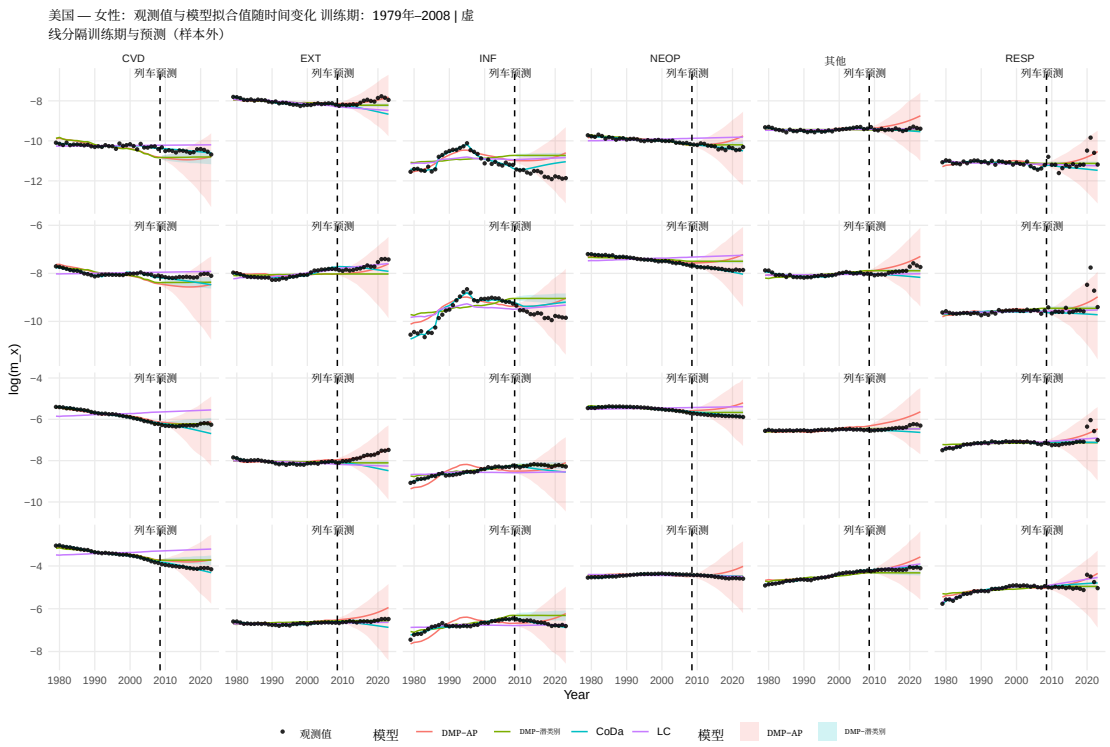

图5：美国女性：选定年龄（20、40、60、80岁）按死因划分的观测值（黑色点）以及拟合/预测对数死亡率。竖直虚线标示样本内时期（1979–2008）的结束和样本外预测的开始。

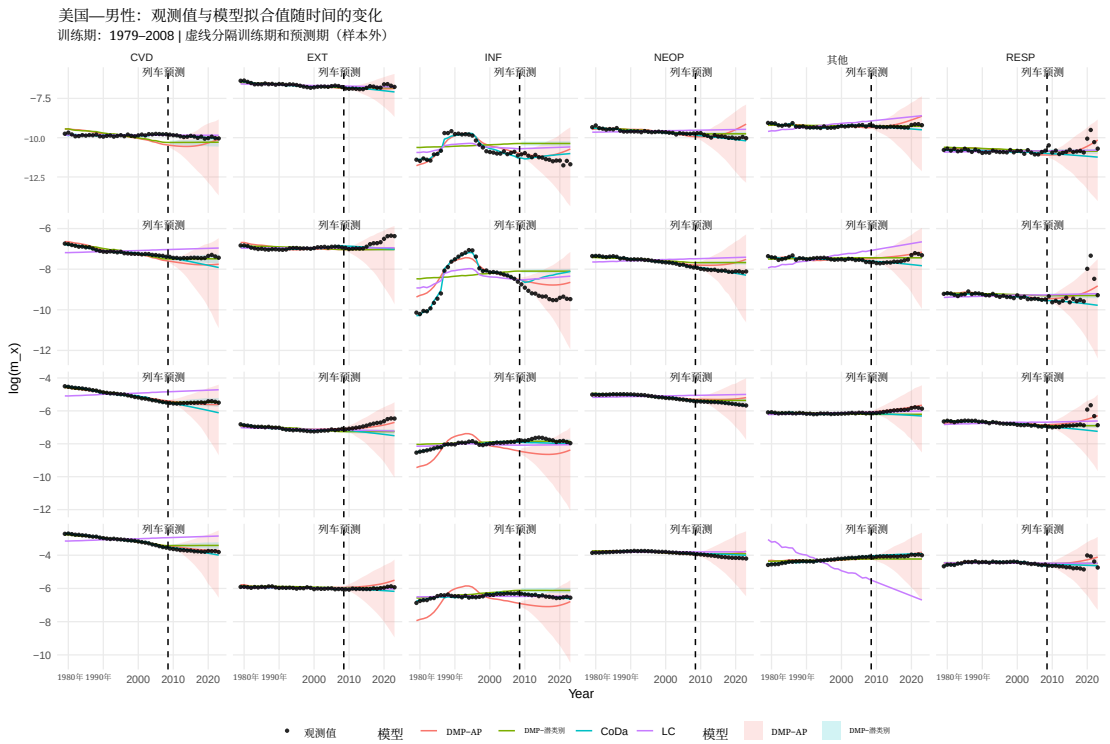

图6：美国男性：选定年龄（20、40、60、80）按死因划分的观测值（黑点）及拟合/预测对数死亡率。竖直虚线标示样本内时期（1979–2008）的结束和样本外预测的开始。

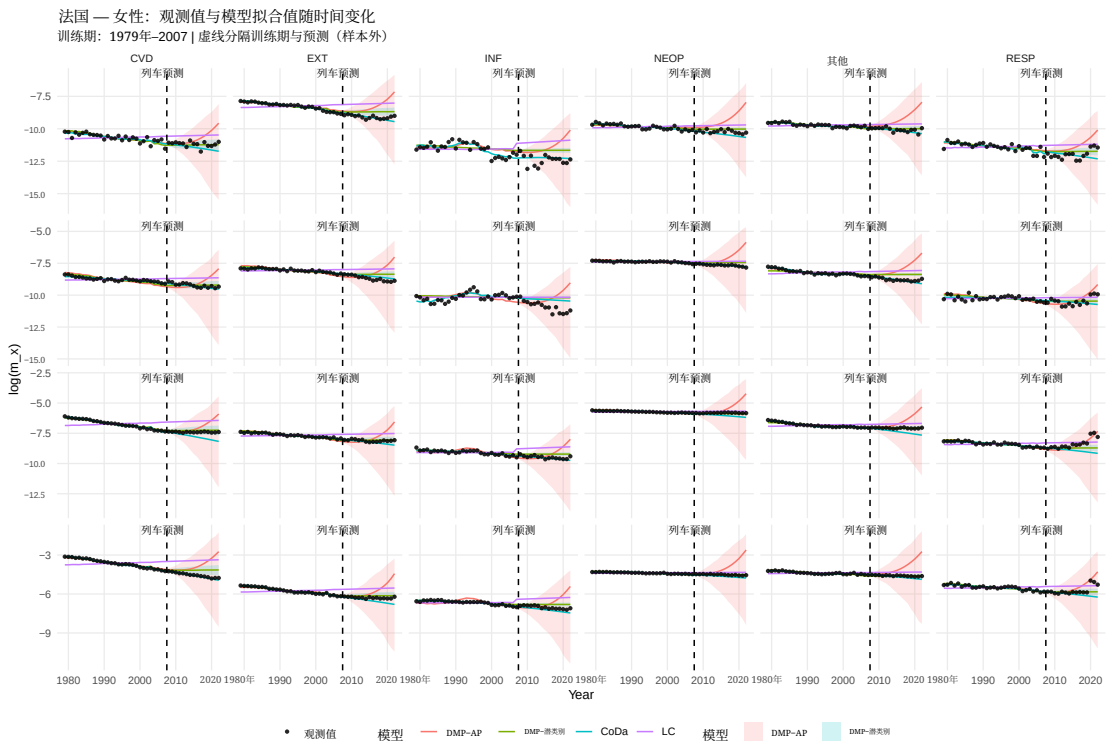

图7：法国女性：选定年龄（20、40、60、80）的分死因对数死亡率观测值（黑点）及拟合/预测值。竖直虚线标示样本内时期（1979–2007）的结束以及样本外预测的开始。

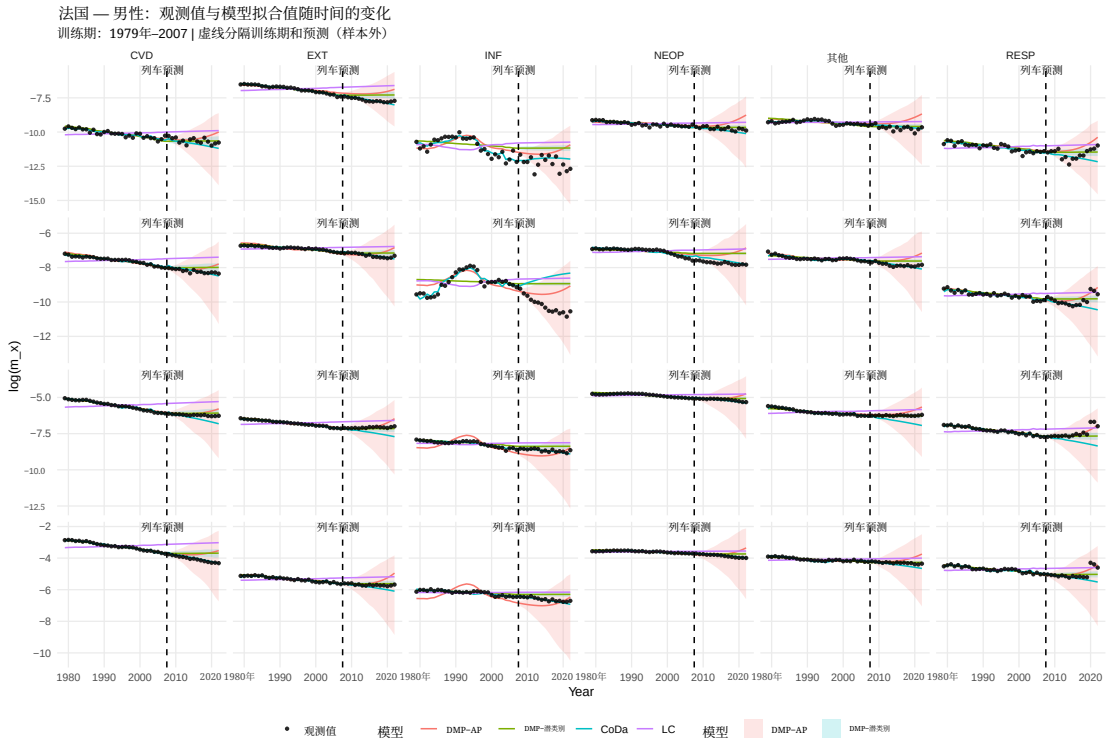

图 8：法国男性：选定年龄（20、40、60、80 岁）各死因的观测对数死亡率（黑点）及拟合值/预测值。竖直虚线标示样本内时期（1979–2007）的结束和样本外预测的开始。

B 基准模型：潜类别与CoDa

我们概述了用于比较的两个基准模型，即拟合死因特异性死亡率的 LC模型，以及拟合生命表死亡分布的 CoDA 方法。我们的实现遵循Kjærgaard 等人（2019年）中的设定。

B.1 逐死因潜类别

令$D_{t,x,c}$ 表示年份 t、年龄组 x、死因 c的观测死亡数，并令$m_{t,x,c}$ 为相应的（分死因）死亡率。LC基准采用标准双线性结构按各死因分别进行拟合

$$\log m_{t,x,c} \;=\; \alpha_{x,c} \;+\; \beta_{x,c}\,\kappa_{t,c} \;+\; \varepsilon_{t,x,c},$$

其中$\alpha_{x,c}$ 为死因 c的平均年龄模式， $\kappa_{t,c}$ 为时间指数，$\beta_{x,c}$ 衡量对$\kappa_{t,c}$ 变化的年龄特异性敏感度（并施加标准可识别性约束，例如$\sum_x \beta_{x,c} = 1$和$\sum_t \kappa_{t,c} \;=\; 0$）。预测通过对$\kappa_{t,c}$ 进行外推（通常采用自回归积分滑动平均（ARIMA）型时间序列模型），随后重构$\widehat{m}_{t,x,c}$而得到。

由于潜类别按死因独立拟合，相干性无法得到保证：一般而言，$\sum_c \widehat{m}_{t,x,c}$不必与独立建模的合计相匹配。因此，我们将潜类别视为一个强但可能非相干的基线（Wilmoth, 1995年）。

B.2 基于生命表死亡人数的 CoDa

CoDa 基准模型对死亡人数的分布进行建模，而非死因别死亡率。从多减因生命表出发，定义全死因生命表死亡数 $d_{t,x}$ （给定一个基数，例如 10^5），并将观测到的死因计数转换为分死因生命表死亡数，方法如下

$$d_{t,x,c} \;=\; d_{t,x}\,\frac{D_{t,x,c}}{D_{t,x}}, \qquad \text{where} \qquad D_{t,x} = \sum_{c=1}^{K} D_{t,x,c}.$$

这样，对于每个年份 t，都会得到一个非负数组 $\{d_{t,x,c}\}$，其中各元素之和为基数，即年龄-死因单元上的构成。

要在一个非受约束的 在受限空间中，CoDa 会应用中心对数比（CLR）变换。如果

我们将年龄-死因单元堆叠为单一索引 $j \in \{1, \ldots, NK\}$，其中 N 是年龄组的数量，K 是死因的数量，并写作 $d_{t,j}$，则中心对数比（CLR）为

$$\operatorname{clr}(d_{t,j}) \;=\; \log\!\left(\frac{d_{t,j}}{g_t}\right), \qquad g_t \;=\; \left(\prod_{j=1}^{NK} d_{t,j}\right)^{1/(NK)}.$$

变换后的序列随后通过奇异值分解进行分解，从而在中心对数比尺度上得到LC型因子分解：

$$\operatorname{clr}(d_{t,j}) \;=\; \alpha_j \;+\; \sum_{p=1}^{P} \beta_{j,p}\, k_{t,p} \;+\; \varepsilon_{t,j}.$$

预测通过对得分向量$k_{t,p}$进行时间序列外推（例如ARIMA）得到，随后经逆中心对数比映射并闭合回到单纯形（从而使预测的死亡分布之和等于基数）。用组成数据运算的记号，逆变换可示意性地写为

$$\widehat{\mathbf{d}}_t \;=\; \operatorname{clr}^{-1}\!\Big(\sum_{p=1}^{P} \boldsymbol{\beta}_p\, \widehat{k}_{t,p}\Big) \;\oplus\; \boldsymbol{\alpha},$$

其中，$\boldsymbol{\alpha}$ 是基准（时间平均）年龄–死因死亡构成，而 $\operatorname{clr}^{-1}(\cdot)$ 包含确保满足总和约束的闭合算子。

C 后验摘要

在表 15 和 16中，我们分别按性别和样本内年份列出了法国和美国贝叶斯 AP 与 LC 模型标量参数的后验汇总。

表15：法国——按性别和样本内起始年（1979年与1989年）划分的标量参数后验摘要（DMP-AP 和 DMP-LC）。

年份	参数	男性							女性						
		Mean	SD	5%	50%	95%	n_{eff}	$\widehat{R}$	Mean	SD	5%	50%	95%	n_{eff}	$\widehat{R}$
DMP-AP															
1979	ν_0	-4.72	0.00	-4.72	-4.72	-4.71	2923	1.00	-5.36	0.00	-5.36	-5.36	-5.36	4465	1.00
	σ_δ	0.67	0.11	0.52	0.65	0.86	667	1.01	0.60	0.10	0.46	0.59	0.78	178	1.00
	σ_π	0.03	0.00	0.03	0.03	0.04	860	1.01	0.05	0.01	0.04	0.05	0.07	204	1.02
	σ_ζ	0.32	0.02	0.28	0.32	0.36	618	1.00	0.28	0.02	0.25	0.28	0.32	211	1.01
	σ_λ	0.02	0.00	0.02	0.02	0.03	508	1.01	0.02	0.00	0.01	0.02	0.02	173	1.03
	ϕ	652.72	21.87	616.82	652.82	688.93	684	1.00	898.07	34.90	844.81	895.57	958.71	167	1.03
DMP-LC															
1979	ν_0	-4.73	0.00	-4.73	-4.73	-4.73	3798	1.00	-5.38	0.00	-5.38	-5.38	-5.38	3106	1.00
	σ_κ	0.65	0.09	0.53	0.64	0.81	522	1.00	0.81	0.10	0.66	0.80	0.99	445	1.01
	σ_ζ	0.31	0.02	0.27	0.30	0.35	438	1.00	0.28	0.02	0.24	0.27	0.31	387	1.01
	ϕ	390.58	11.25	372.88	390.25	409.77	597	1.01	667.74	23.44	629.35	667.30	707.14	268	1.01
DMP-AP															
1989	ν_0	-4.83	0.00	-4.83	-4.83	-4.83	4608	1.00	-5.48	0.00	-5.48	-5.48	-5.48	5609	1.00
	σ_δ	0.66	0.11	0.51	0.64	0.87	631	1.00	0.60	0.11	0.46	0.58	0.80	366	1.02
	σ_π	0.04	0.01	0.03	0.04	0.05	850	1.01	0.06	0.01	0.05	0.06	0.08	412	1.00
	σ_ζ	0.34	0.03	0.30	0.34	0.39	711	1.00	0.31	0.02	0.28	0.31	0.35	496	1.00
	σ_λ	0.03	0.00	0.03	0.03	0.04	545	1.00	0.03	0.00	0.02	0.03	0.03	379	1.01
	ϕ	839.40	36.75	781.70	838.02	902.39	595	1.00	2032.95	130.66	1832.24	2027.17	2256.96	433	1.01
DMP-LC															
1989	ν_0	-4.84	0.00	-4.84	-4.84	-4.84	2400	1.00	-5.49	0.00	-5.49	-5.49	-5.49	4425	1.00
	σ_κ	0.75	0.12	0.58	0.74	0.97	285	1.01	0.89	0.13	0.71	0.88	1.12	447	1.00
	σ_ζ	0.34	0.03	0.30	0.34	0.39	229	1.00	0.31	0.02	0.27	0.30	0.35	350	1.01
	ϕ	634.99	24.50	593.58	635.33	675.55	194	1.01	1316.73	74.93	1195.37	1316.18	1441.40	610	1.00

表16：美国——标量参数的后验摘要（DMP-AP 和 DMP-LC），按性别和样本内起始年（1979年与1989年）划分。

年份	参数	男性							女性						
		Mean	SD	5%	50%	95%	n_{eff}	$\widehat{R}$	Mean	SD	5%	50%	95%	n_{eff}	$\widehat{R}$
DMP-AP															
1979	ν_0	-4.65	0.00	-4.65	-4.65	-4.65	2674	1.00	-5.18	0.00	-5.18	-5.18	-5.18	2990	1.00
	σ_δ	0.70	0.12	0.54	0.68	0.91	323	1.00	0.63	0.11	0.48	0.62	0.84	195	1.02
	σ_π	0.02	0.00	0.01	0.02	0.02	210	1.00	0.02	0.00	0.02	0.02	0.03	264	1.00
	σ_ζ	0.34	0.03	0.30	0.34	0.39	269	1.03	0.32	0.03	0.28	0.32	0.37	225	1.00
	σ_λ	0.03	0.00	0.02	0.03	0.04	249	1.00	0.01	0.00	0.01	0.01	0.02	96	1.04
	ϕ	483.17	12.47	462.16	483.03	503.70	217	1.00	591.35	16.80	562.09	591.38	618.97	242	1.01
DMP-LC															
1979	ν_0	-4.66	0.00	-4.66	-4.66	-4.66	3172	1.00	-5.18	0.00	-5.18	-5.18	-5.18	2328	1.00
	σ_κ	0.37	0.05	0.30	0.37	0.46	273	1.01	0.31	0.04	0.25	0.31	0.39	340	1.00
	σ_ζ	0.33	0.02	0.29	0.33	0.37	186	1.00	0.32	0.02	0.28	0.31	0.36	290	1.02
	ϕ	229.22	5.85	219.36	229.40	238.45	298	1.01	416.47	10.76	398.33	416.54	434.21	275	1.01
DMP-AP															
1989	ν_0	-4.72	0.00	-4.72	-4.72	-4.72	2549	1.00	-5.22	0.00	-5.22	-5.22	-5.22	2984	1.00
	σ_δ	0.71	0.12	0.54	0.70	0.94	387	1.01	0.64	0.11	0.49	0.62	0.85	258	1.03
	σ_π	0.02	0.00	0.01	0.02	0.03	425	1.00	0.02	0.00	0.02	0.02	0.03	378	1.01
	σ_ζ	0.35	0.03	0.31	0.35	0.40	254	1.00	0.32	0.02	0.29	0.32	0.36	346	1.02
	σ_λ	0.03	0.00	0.02	0.03	0.03	286	1.02	0.02	0.00	0.01	0.02	0.02	231	1.01
	ϕ	615.65	21.02	579.38	615.89	649.42	510	1.00	928.29	31.49	876.75	928.39	980.45	426	1.01
DMP-潜类别															
1989	ν_0	-4.72	0.00	-4.72	-4.72	-4.72	2842	1.00	-5.23	0.00	-5.23	-5.23	-5.22	3424	1.00
	σ_κ	0.44	0.08	0.34	0.43	0.59	491	1.01	0.31	0.05	0.24	0.30	0.40	232	1.02
	σ_ζ	0.35	0.03	0.31	0.35	0.39	467	1.00	0.32	0.02	0.29	0.32	0.36	291	1.00
	ϕ	545.39	18.20	516.78	545.21	574.64	472	1.00	805.97	28.19	759.64	805.95	851.91	273	1.00