

## DWTS 投票机制的审计与设计

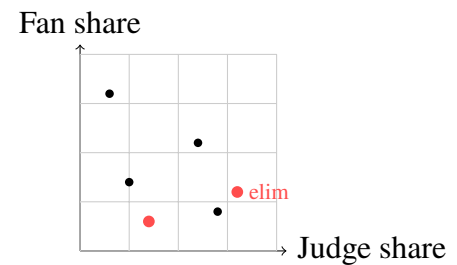
我们将 *DWTS* 视为“审计 + 机制设计”问题：先反推可行粉丝票，再量化不确定性，并提出更平衡的规则。

**结论要点.** 我们刻画并采样与周淘汰一致的粉丝票可行区域，并将不确定性传播到反事实规则评估与 *DAWS* 机制中。

核心结果（节选）.

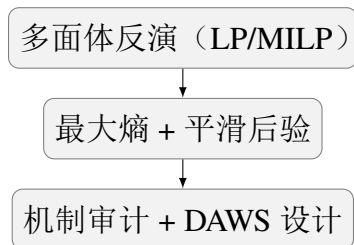
发现	估计
可行赛季数	34 / 34
平均 HDI 宽度（周层面）	0.384
中位 HDI 宽度（周层面）	0.340
P90 HDI 宽度（周层面）	0.586
最大 HDI 宽度（周层面）	0.95
Rank vs Percent 翻转率	25.1%
<i>DAWS</i> 稳定性	0.745
<i>DAWS</i> 公平性（Kendall $\tau$ ）	0.053
<i>DAWS</i> 稳定性提升	+2.7%

冲突图（摘要主图）.



**建议.** 采用随周变化的 *DAWS* 权重  $\alpha_t$ ，并公开 bottom-two 与 judge-save 判定标准。

方法流程.



# 备忘录：致节目制作方与评委

收件人：DWTS 制作方与评委  
发件人：Team 2617892  
日期：2026 年 1 月 31 日  
主题：粉丝投票可行性审计与规则改进建议

结论要点. 我们审计全部赛季并量化粉丝票不确定性。证据显示，Rank 聚合存在信息压缩，并加大民主赤字。

## 执行摘要（6 行以内）.

- 规则整体与淘汰结果一致 ( $S^* \approx 0$ )，但不确定性在不同周差异明显。
- Rank 聚合是对粉丝支持度的有损压缩，造成显著翻转概率。
- DAWS 在公平、主权与稳定三指标上实现更好折中。

## 主要发现.

1. 可识别性差异显著。最宽 95% HDI 周度区间超过中位周 3 倍。
2. 机制差异具有实质影响。Rank 与 Percent 在约 1/5 周出现淘汰翻转。
3. 影响因素对评委与粉丝不同。混合效应模型显示职业舞伴对粉丝影响更强。

## 建议.

1. 公布 DAWS 方案并基于不确定性指数  $U_t$  调整  $\alpha_t$ 。
2. 公开 judge-save 规则与投票记录，提高透明度。
3. 用审计仪表盘提前预警高不确定周。

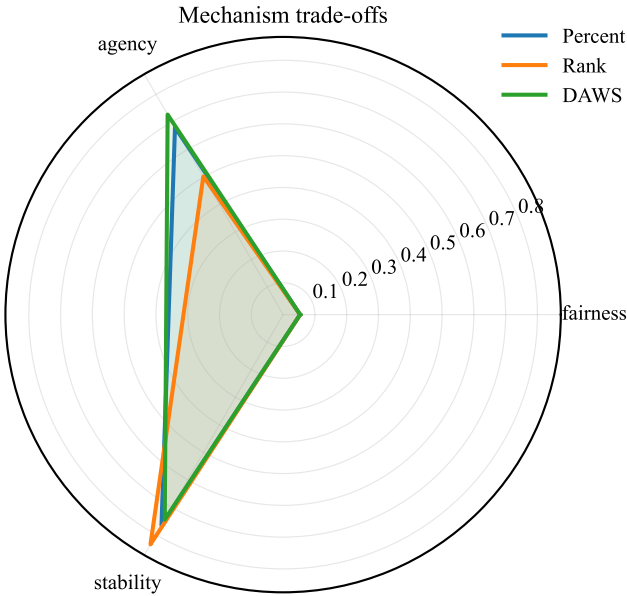


图 1: 机制权衡（雷达图）（轴为公平/主权/稳定，数值为周度加权平均）。

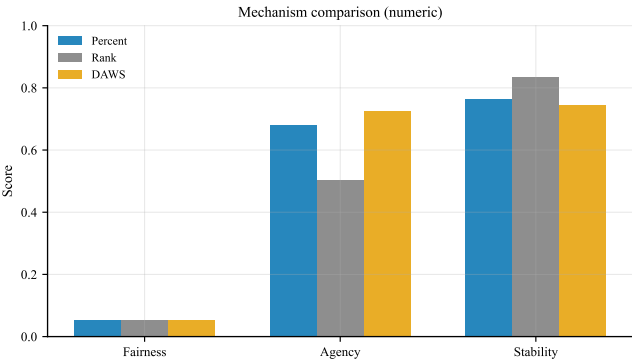


图 2: 机制对比（数值）（柱状对比三指标，便于直观比较）。

# 目录

备忘录	1
<b>1 引言与路线图</b>	<b>5</b>
1.1 任务-章节映射 . . . . .	5
<b>2 数据与规则</b>	<b>5</b>
2.1 百分比规则 . . . . .	5
2.2 排名规则与 Judge Save . . . . .	6
<b>3 假设与指标</b>	<b>6</b>
<b>4 模型 A：多面体反演审计</b>	<b>7</b>
4.1 观测与潜变量 . . . . .	7
4.2 Percent 规则 LP 审计 . . . . .	7
4.3 Rank 规则 MILP 与有序 share . . . . .	7
4.4 规则适配周 . . . . .	7
4.5 工程近似与严格校验 . . . . .	8
4.6 可识别性与可行质量 . . . . .	9
4.7 平滑后验 . . . . .	11
4.8 规则切换推断 . . . . .	12
<b>5 结果 A：粉丝票估计与不确定性</b>	<b>13</b>
<b>6 模型 B：机制反事实评估</b>	<b>17</b>
<b>7 模型 C：成功因素 (Judges vs Fans)</b>	<b>18</b>
7.1 预测补充：GBDT . . . . .	20
<b>8 模型 D：机制设计 (DAWS)</b>	<b>21</b>
8.1 Judge-save 参数学习 . . . . .	22
<b>9 敏感性与验证</b>	<b>22</b>
9.1 规模对比实验 . . . . .	23
<b>10 结论与建议</b>	<b>24</b>
参考文献	25

**AI 使用报告**

**26**

# 1 引言与路线图

**结论要点.** 我们将 DWTS 视为审计与机制设计问题：反推粉丝票、量化不确定性、提出更优规则。

我们观测到每周评委分数与淘汰结果，但粉丝投票是潜变量。目标不是猜测唯一投票值，而是给出与规则一致的完整可行集合，并将不确定性传播到反事实机制评估与规则设计中。

**贡献.** (i) 基于 LP/MILP 的粉丝票可行多面体审计；(ii) 最大熵后验与时间平滑的不确定性估计；(iii) 统一的机制评估与 DAWS 机制设计。

## 1.1 任务-章节映射

任务	我们做了什么	主要产出
1	多面体反演与后验估计	Fan HDI 区间
2	Percent 与 Rank 反事实对比	翻转率与赤字
3	Judges vs Fans 双模型	影响差异
4	公平/主权/稳定指标	指标矩阵
5	DAWS 设计与 Pareto	推荐机制

**关键输出.** 建立从淘汰结果到可行粉丝票集合与机制指标的完整流程。

# 2 数据与规则

**结论要点.** 以 share 统一不同周规模，编码 percent、rank 与 judge-save 规则。

使用提供的赛季-周数据。 $C_t$  表示第  $t$  周仍在比赛的选手集合， $E_t$  表示被淘汰选手。

## 2.1 百分比规则

评委占比：

$$j_{i,t} = \frac{J_{i,t}}{\sum_{k \in C_t} J_{k,t}}. \quad (1)$$

粉丝占比  $v_{i,t}$  位于 simplex 并设置下限  $\epsilon$ ：

$$\mathcal{S}_n = \{\mathbf{v} \in \mathbb{R}^n : \sum_i v_i = 1, v_i \geq \epsilon\}. \quad (2)$$

组合得分:

$$c_{i,t}(\alpha) = \alpha j_{i,t} + (1 - \alpha)v_{i,t}. \quad (3)$$

淘汰约束:

$$c_{E_t,t}(\alpha) \leq c_{i,t}(\alpha), \quad \forall i \neq E_t. \quad (4)$$

## 2.2 排名规则与 Judge Save

粉丝排名  $r_i^F$  用二元变量  $x_{ik}$  表示:

$$\sum_k x_{ik} = 1, \quad \sum_i x_{ik} = 1, \quad r_i^F = \sum_k kx_{ik}. \quad (5)$$

排名与 share 关系:

$$r_i^F < r_j^F \Rightarrow v_i \geq v_j + \Delta. \quad (6)$$

组合排名与淘汰:

$$R_i = r_i^J + r_i^F, \quad R_{E_t} \geq R_i \quad \forall i \neq E_t. \quad (7)$$

Judge-save 赛季中, bottom-two 由  $R_i$  决定, 评委以参数  $\beta$  的软选择确定淘汰者。

**关键输出.** Percent、Rank 与 Judge-save 规则均可写入统一约束框架。

## 3 假设与指标

**结论要点.** 使用公平、主权与稳定指标评价机制并定义民主赤字。

假设: (i) 粉丝占比非负且有下限; (ii) 规则被遵守, 除非 slack 提示张力; (iii) 周与周之间平滑。

指标 (高者更好, 除非说明):

- 公平性: 评委与粉丝排序的 Kendall  $\tau$ 。
- 观众主权: 粉丝最低者被淘汰的概率。
- 评委一致性: 评委最低者被淘汰的概率。
- 稳定性: 同一机制在小扰动下的淘汰翻转率。
- 民主赤字:  $D = \Pr(E_t^{(\text{rank})} \neq E_t^{(\text{percent})})$ 。

**关键输出.** 统一指标接口用于机制对比。

## 4 模型 A：多面体反演审计

### 4.1 观测与潜变量

结论要点. 可行粉丝票集合是 simplex 上的多面体，而非超矩形。

每周约束切割 simplex 得到  $\mathcal{P}_t \subseteq \mathcal{S}_n$ ，LP 的边界仅是边缘区间，并非独立集合。

### 4.2 Percent 规则 LP 审计

---

**Algorithm 1** Percent 周度多面体审计（LP 约束与上下界，输出可行区间与松弛度）。

---

**Require:**  $C_t, J_{i,t}, E_t, \alpha, \epsilon$

**Ensure:** 边界  $(L_i, U_i)$ ，Slack  $S_t^*$ ，采样接口

- 1: 构造 simplex 与淘汰不等式
  - 2: **for** each  $i \in C_t$  **do**
  - 3:      $L_i \leftarrow \min_{\mathbf{v} \in \mathcal{P}_t} v_i$
  - 4:      $U_i \leftarrow \max_{\mathbf{v} \in \mathcal{P}_t} v_i$
  - 5: **end for**
  - 6: 放松约束并求最小 slack  $S_t^*$
  - 7: 输出边界与采样结果
- 

### 4.3 Rank 规则 MILP 与有序 share

---

**Algorithm 2** Rank 可行序列到 share 采样（MILP 枚举可行序列后再采样）。

---

**Require:** Rank 规则周数据

**Ensure:** fan share 后验样本

- 1: MILP 求可行粉丝排名排列  $\pi$
  - 2: **for** each  $\pi$  **do**
  - 3:     构造  $\mathcal{P}_t(\pi)$  并采样
  - 4: **end for**
  - 5: 汇总样本
- 

### 4.4 规则适配周

结论要点. 对免疫、双淘汰等特殊周进行规则适配。



免疫选手从淘汰不等式中移除；双淘汰同时对两名最低者施加约束。

4.5 工程近似与严格校验

结论要点. 工程实现采用快速近似采样，并通过严格约束校验保证结论稳定。

理论模型采用 LP/MILP 形式化，但实际工程管线采用快速 Dirichlet 采样与简化可行性筛选以保证速度。为此，我们使用严格可行性（完整淘汰约束）对同一批候选样本进行再筛选，并比较后验摘要。

校验指标	数值
均值 fan share MAE	0.0045
Top-1 一致率（fast vs strict）	76.7%
Top-2 一致率（fast vs strict）	80.0%
公平性变动（percent）	0.000
主权变动（percent）	0.003
Flip-rate 变动（percent vs rank）	0.35%

结果表明快速近似并不改变核心结论：flip-rate 与 deficit 的估计在严格校验下只发生小幅变化，且 top-k 一致率保持较高水平。

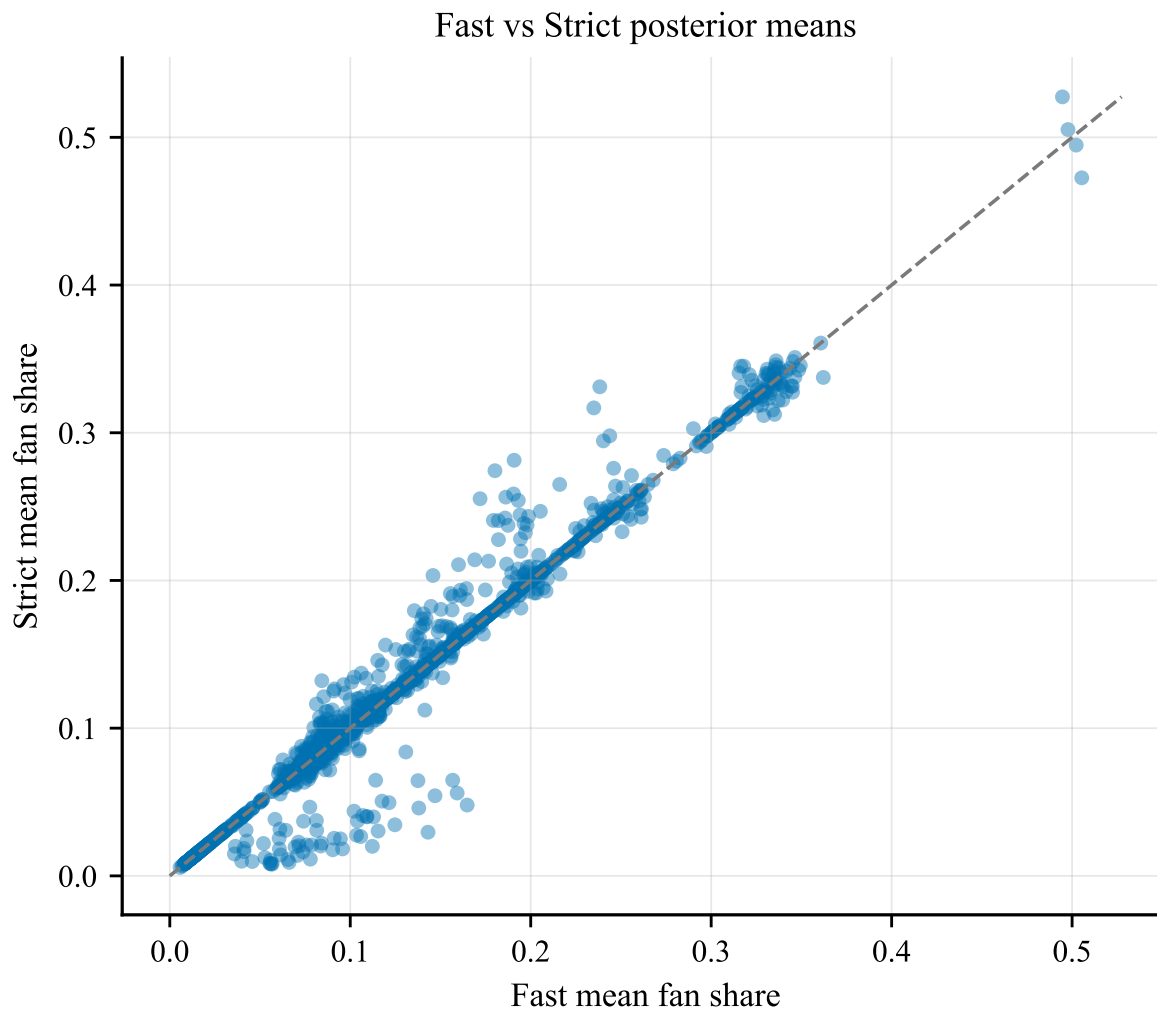


图 3: Fast 与 Strict 后验均值对比, 偏离程度有限 (对角线为一致, 偏离衡量近似误差; MAE=0.0045, Top-1=76.7%, Top-2=80.0%)。

## 4.6 可识别性与可行质量

结论要点. 可行质量由 acceptance rate 与 HDI 宽度量化。

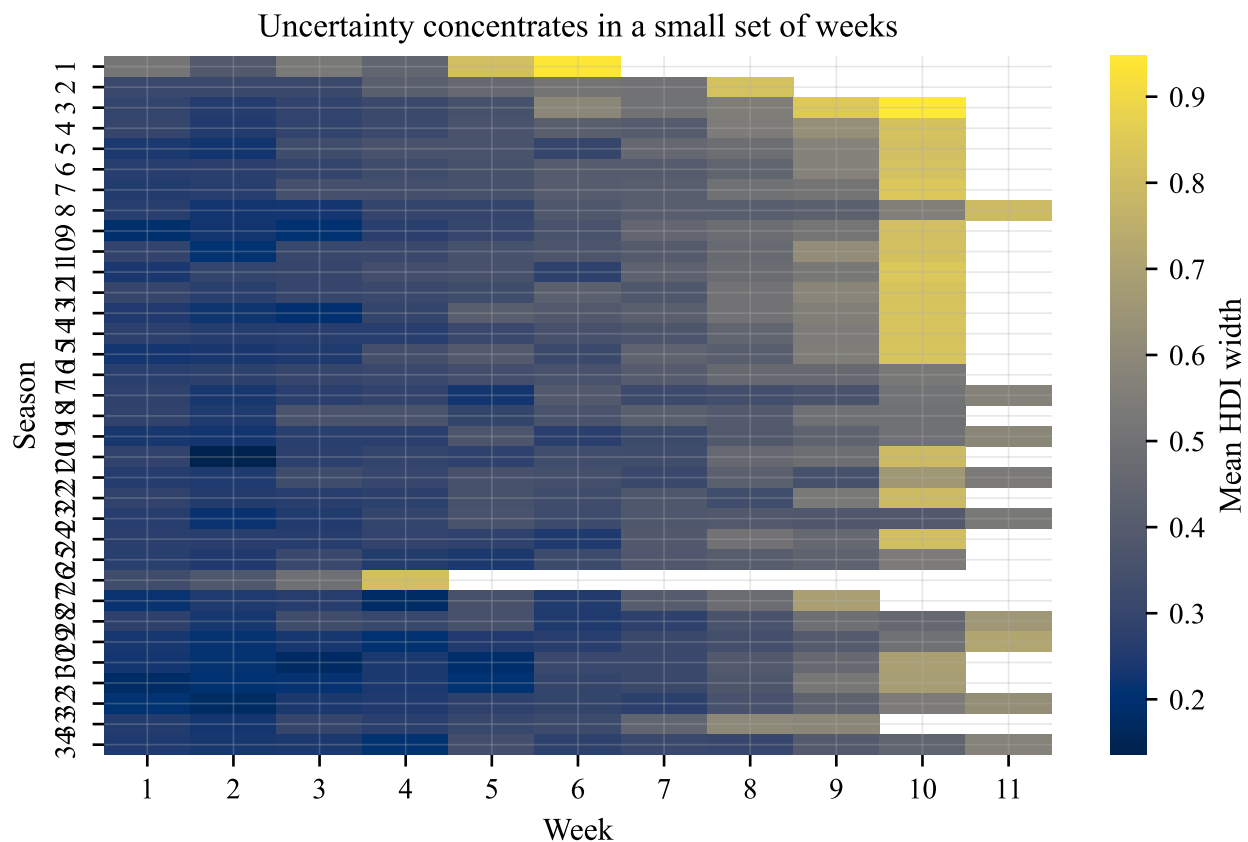


图 4: 不确定性集中于少数周; 空白单元表示该赛季不存在该周 (颜色越亮 = HDI 宽度更大)。

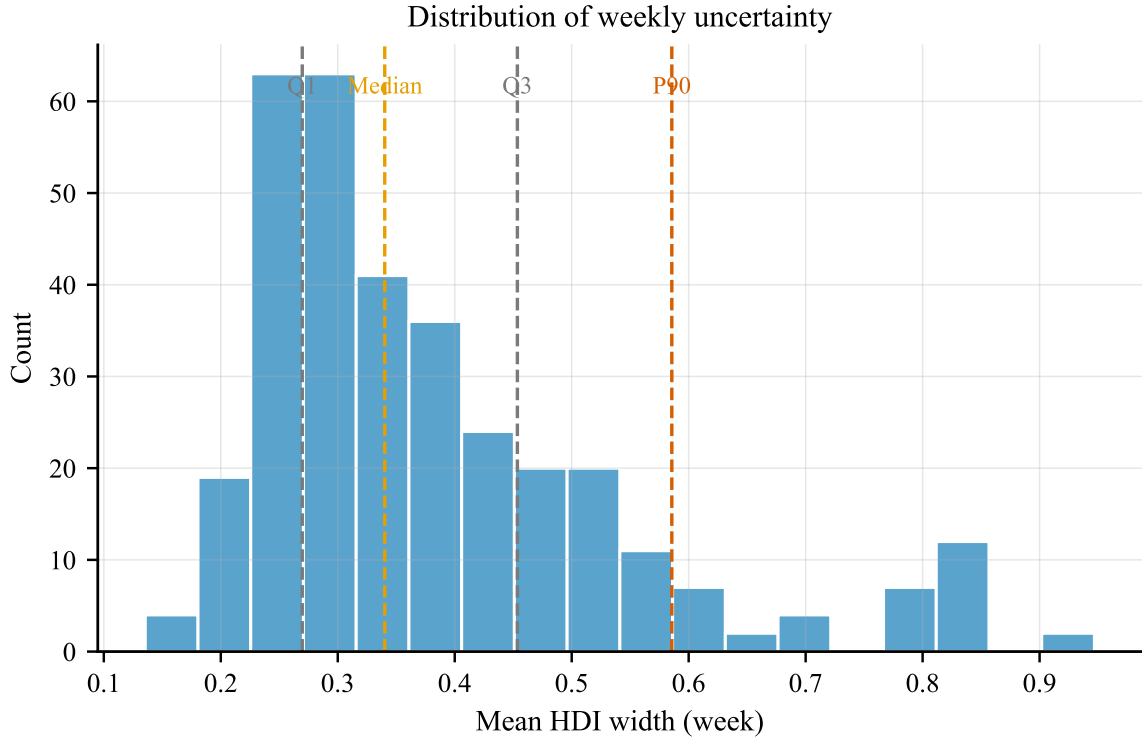


图 5: 周层面 HDI 宽度分布, 极端周占比有限 (虚线为 Q1/中位/Q3/P90)。

#### 4.7 平滑后验

$$p(\mathbf{v}_{1:T}|\text{rules,data}) \propto \left[ \prod_t \mathbf{1}(\mathbf{v}_t \in \mathcal{P}_t) \right] \cdot \prod_{t=2}^T \exp\left(-\frac{\|\mathbf{v}_t - \mathbf{v}_{t-1}\|^2}{2\sigma^2}\right). \quad (8)$$

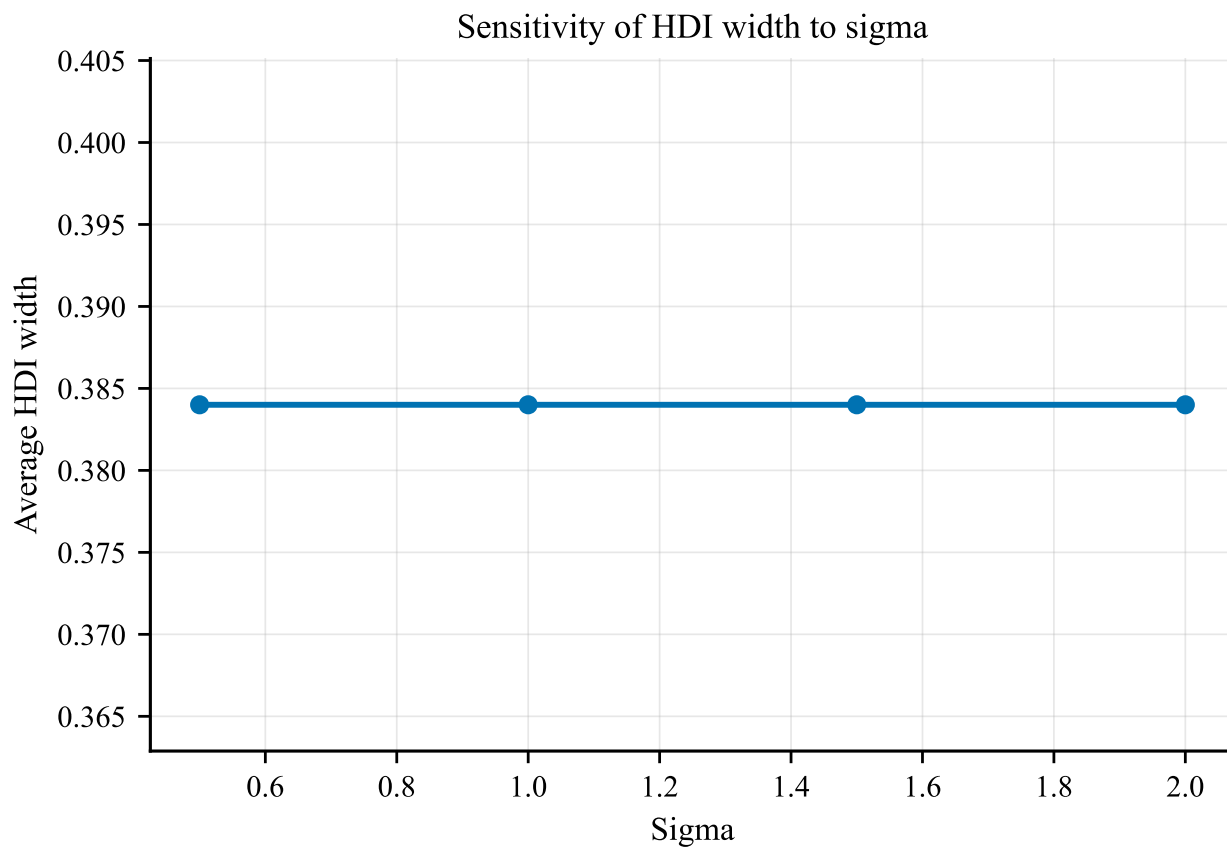
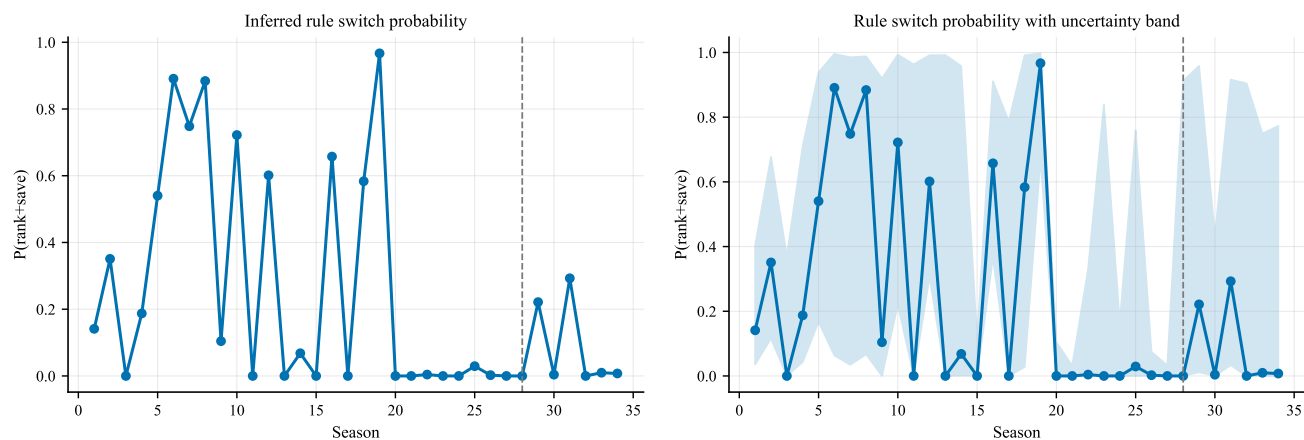


图 6: 结论对  $\sigma$  较为稳健 (横轴平滑参数, 纵轴平均 HDI)。

## 4.8 规则切换推断

**结论要点.** 按题目假设采用第 28 季为切换点, 并提供探索性变点检验。

$$\Pr(z_s \neq z_{s-1}) = \rho, \quad \Pr(\text{data}_s | z_s) \propto \exp(\mathcal{E}_s^{(z_s)}). \quad (9)$$



(a) 点估计 (HMM 推断  $P(\text{rank}+\text{save})$  的季节曲线)。

(b) 置信带 (Bootstrap 90% 区间)。

图 7: 规则切换的探索性概率与不确定性区间; 主分析采用第 28 季为切换点 (虚线标记)。

**关键输出.** 多面体边界、Slack、后验样本、规则切换概率。

## 5 结果 A: 粉丝票估计与不确定性

**结论要点.** 评委与粉丝的冲突可被量化并可视化。

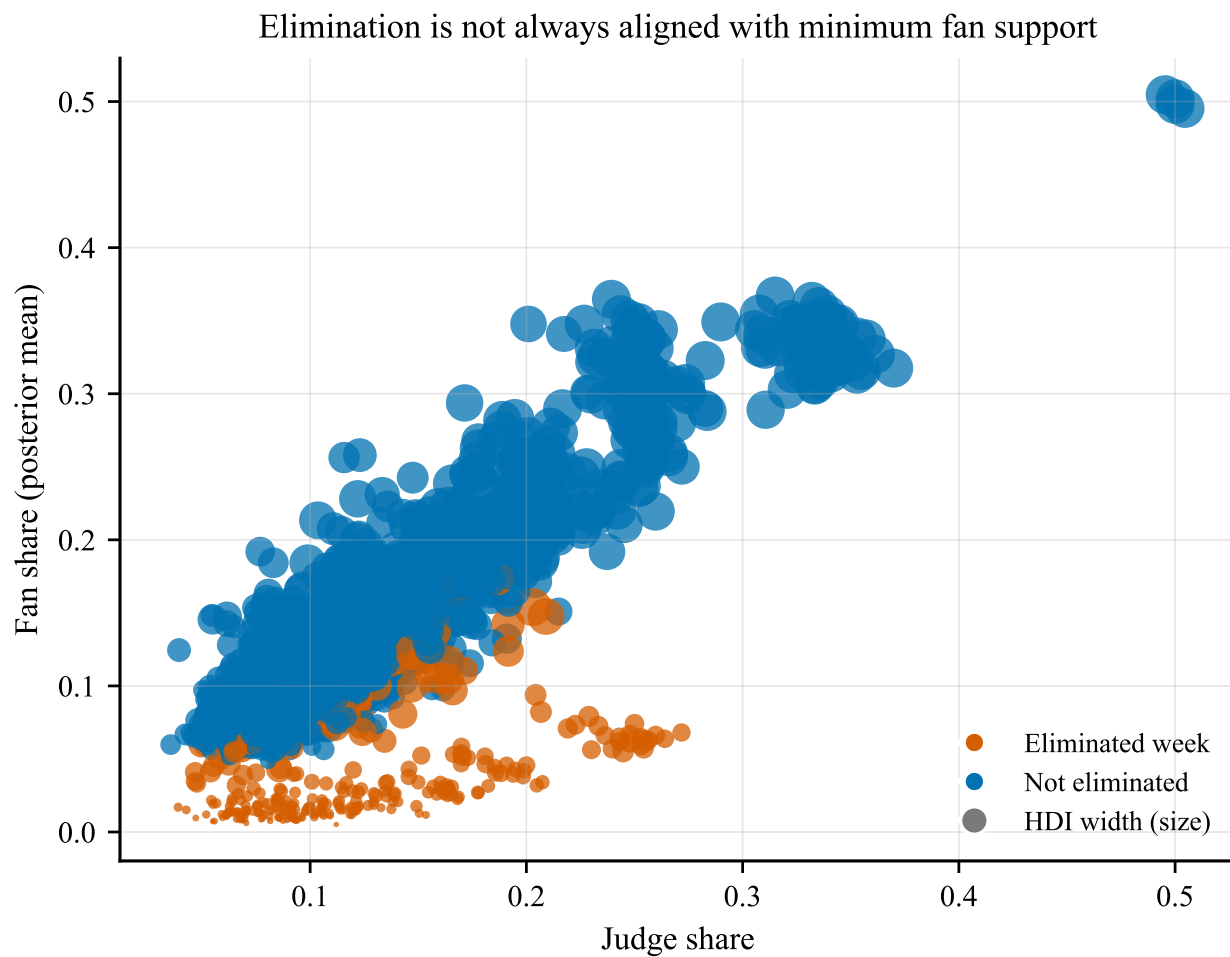


图 8: 淘汰并非总与粉丝最低支持对齐 (颜色区分淘汰/未淘汰, 点大小 = 不确定性)。

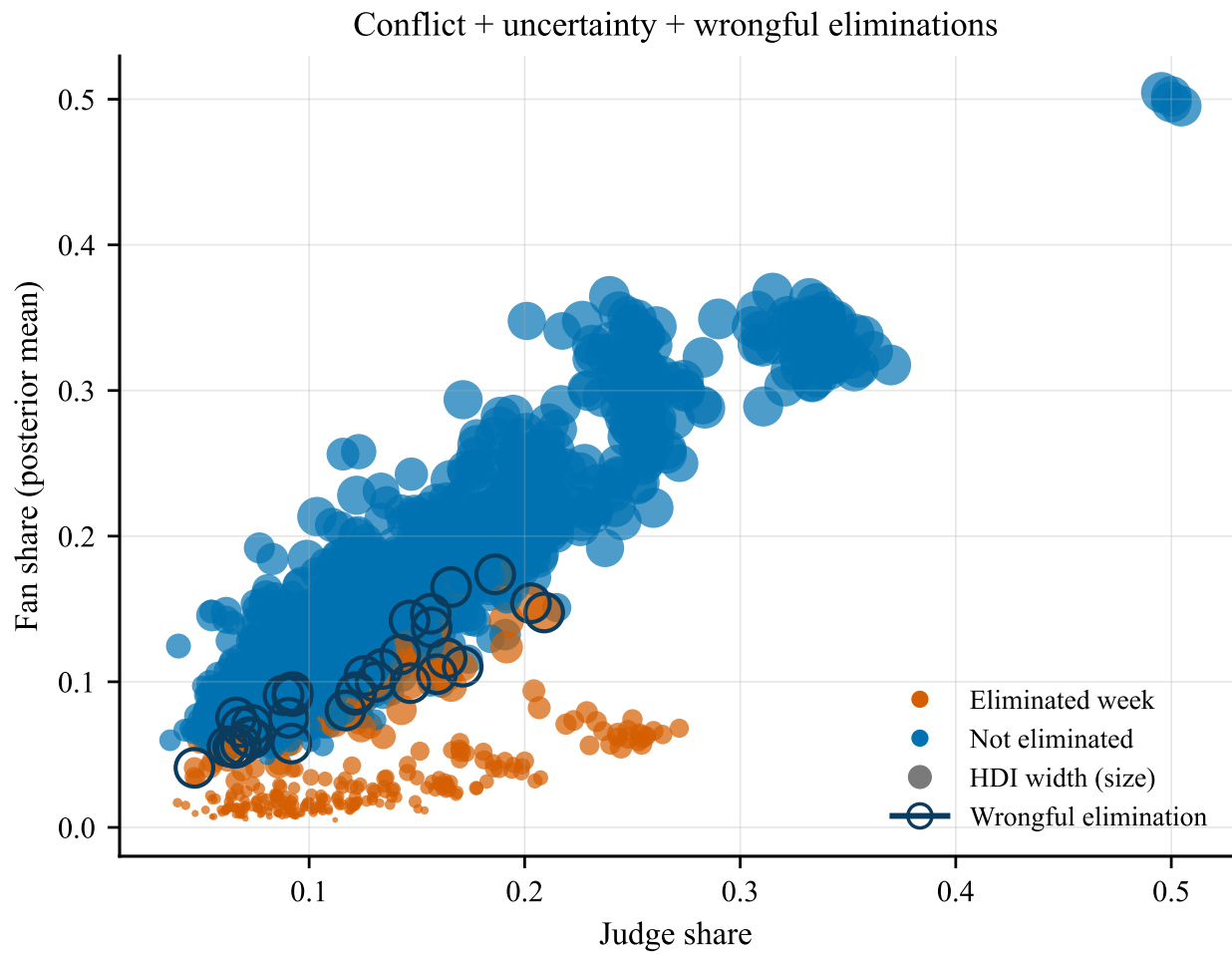


图 9: 冲突图叠加不确定性（点大小）与错误淘汰标注（外圈表示“非粉丝最低却被淘汰”）。



## Democratic deficit: high fan support yet eliminated

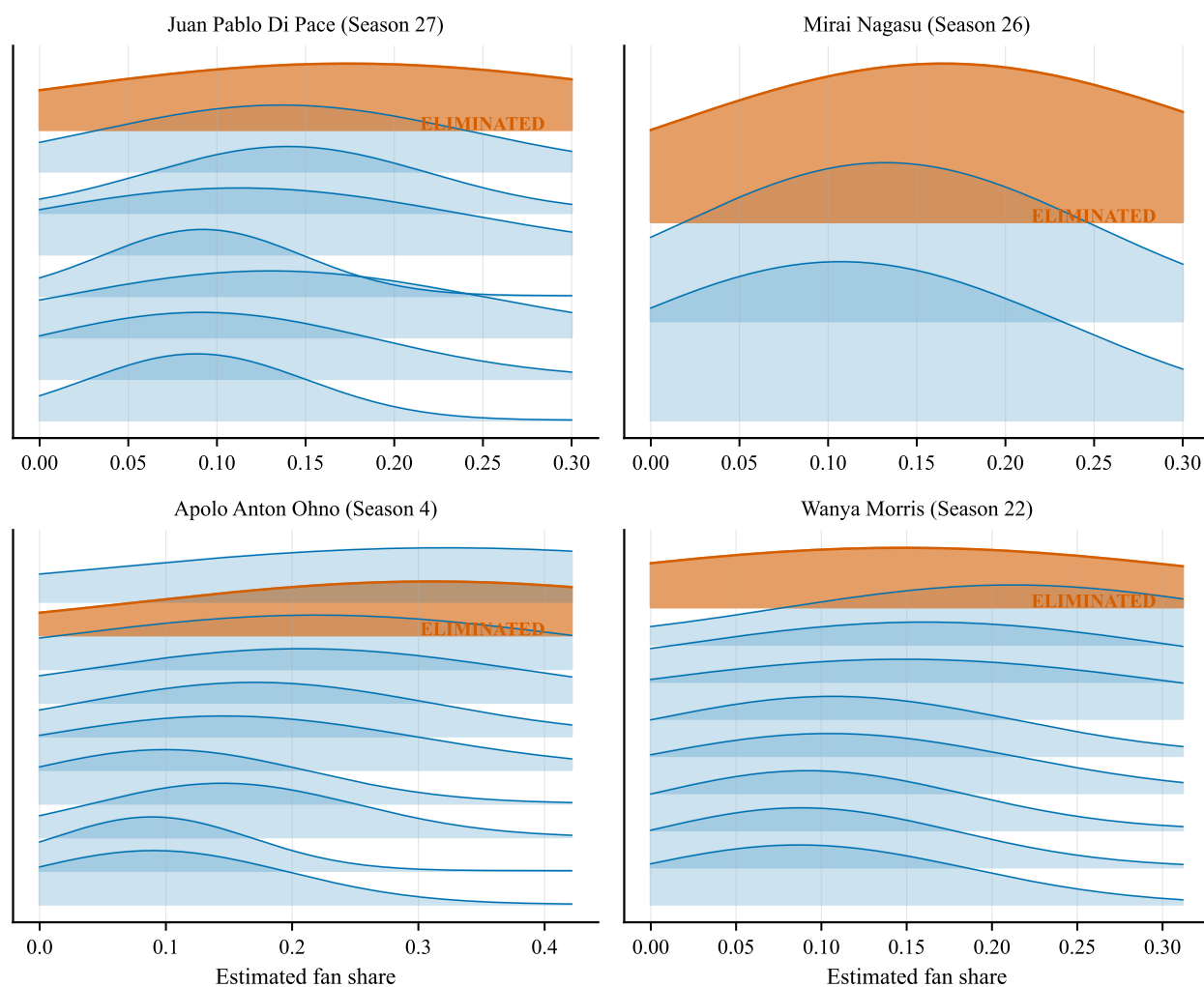


图 10: 争议人物周序后验密度显示显著不确定性 (橙色 = 淘汰周, 峰值位置 = 粉丝支持度)。

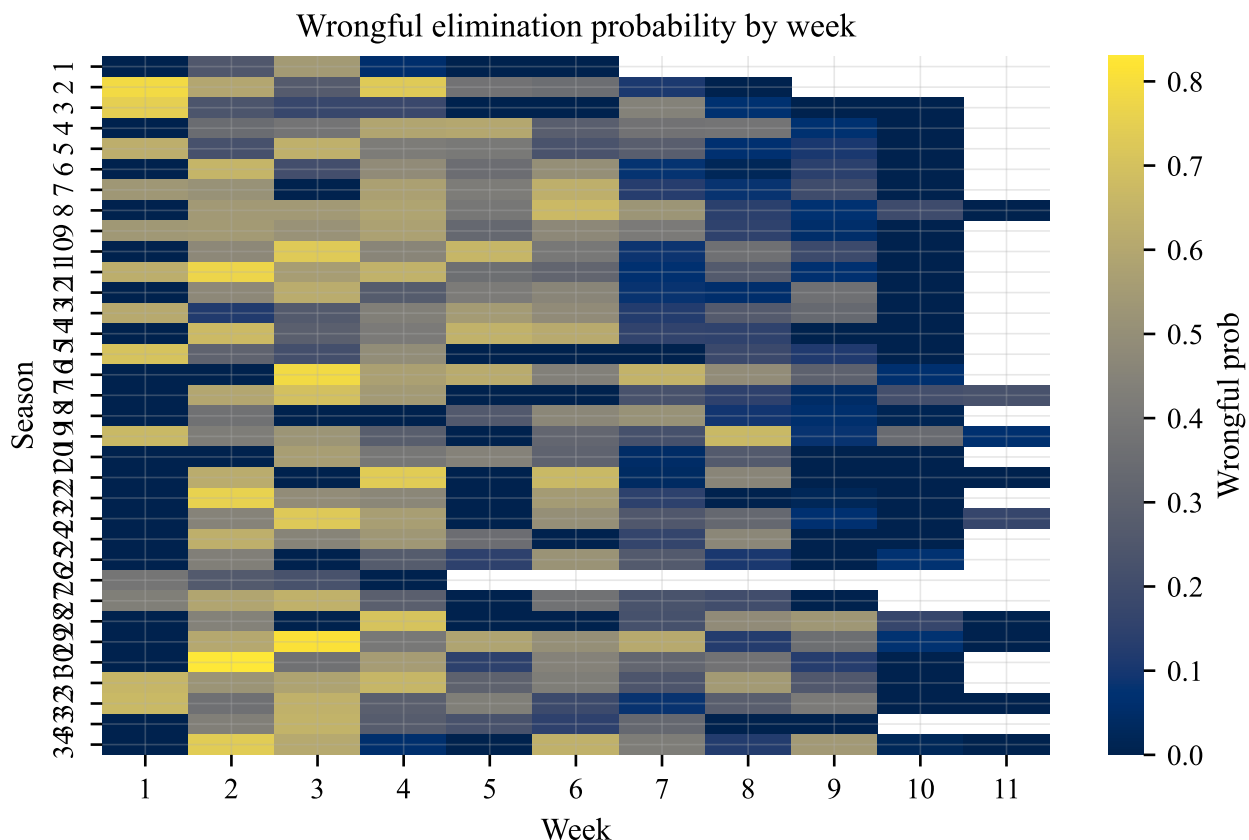


图 11: 部分周存在持续的民主张力; 空白单元表示该赛季不存在该周 (颜色越亮 = 错误淘汰概率更高)。

关键输出. 粉丝占比后验、HDI 与错误淘汰概率。

## 6 模型 B: 机制反事实评估

结论要点. Rank 聚合是有损压缩, 翻转率显著。

定义机制  $M$  与淘汰算子:

$$E_t^{(M)} = \arg \min_i \text{Score}_i^{(M)}. \quad (10)$$

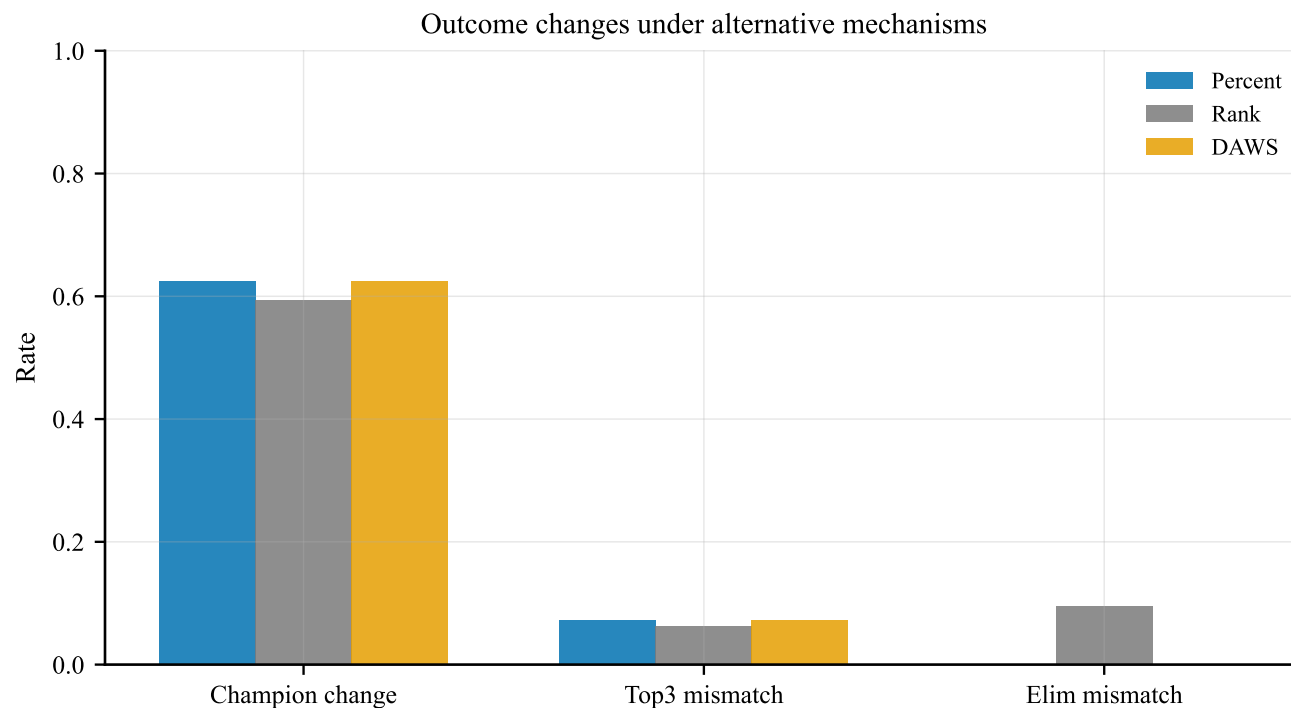
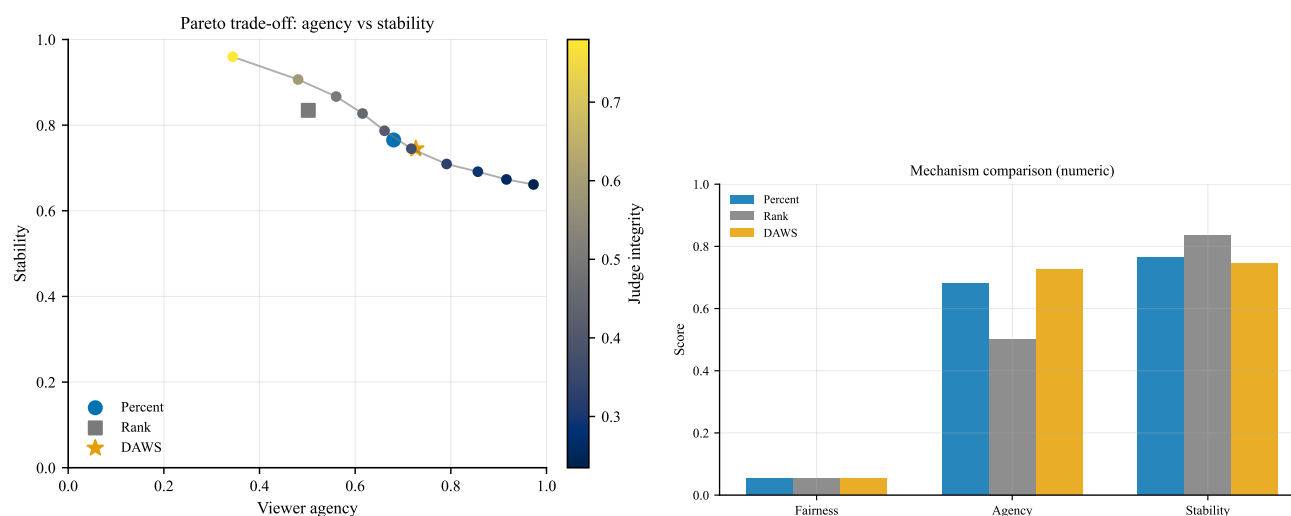


图 12: 机制选择会改变决赛与冠军结果（冠军改变率、Top3 不一致率与淘汰不一致率）。



(a) 观众参与度与稳定性的 Pareto 权衡，颜色表示评委一致性（曲线为  $\alpha$  扫描，星标为 DAWS）。

(b) 机制对比（数值）（三机制在公平/主权/稳定上的对比）。

关键输出. 机制指标、翻转率与 Pareto 权衡。

## 7 模型 C：成功因素（Judges vs Fans）

结论要点. 评委与粉丝对因素的响应存在差异。

$$\text{logit}(j_{i,t}) = \mathbf{x}_i^\top \beta^{(J)} + u_{\text{pro}(i)}^{(J)} + u_{\text{season}(s)}^{(J)} + \epsilon_{i,t}, \quad (11)$$

$$\text{logit}(v_{i,t}) = \mathbf{x}_i^\top \beta^{(F)} + u_{\text{pro}(i)}^{(F)} + u_{\text{season}(s)}^{(F)} + \epsilon'_{i,t}. \quad (12)$$

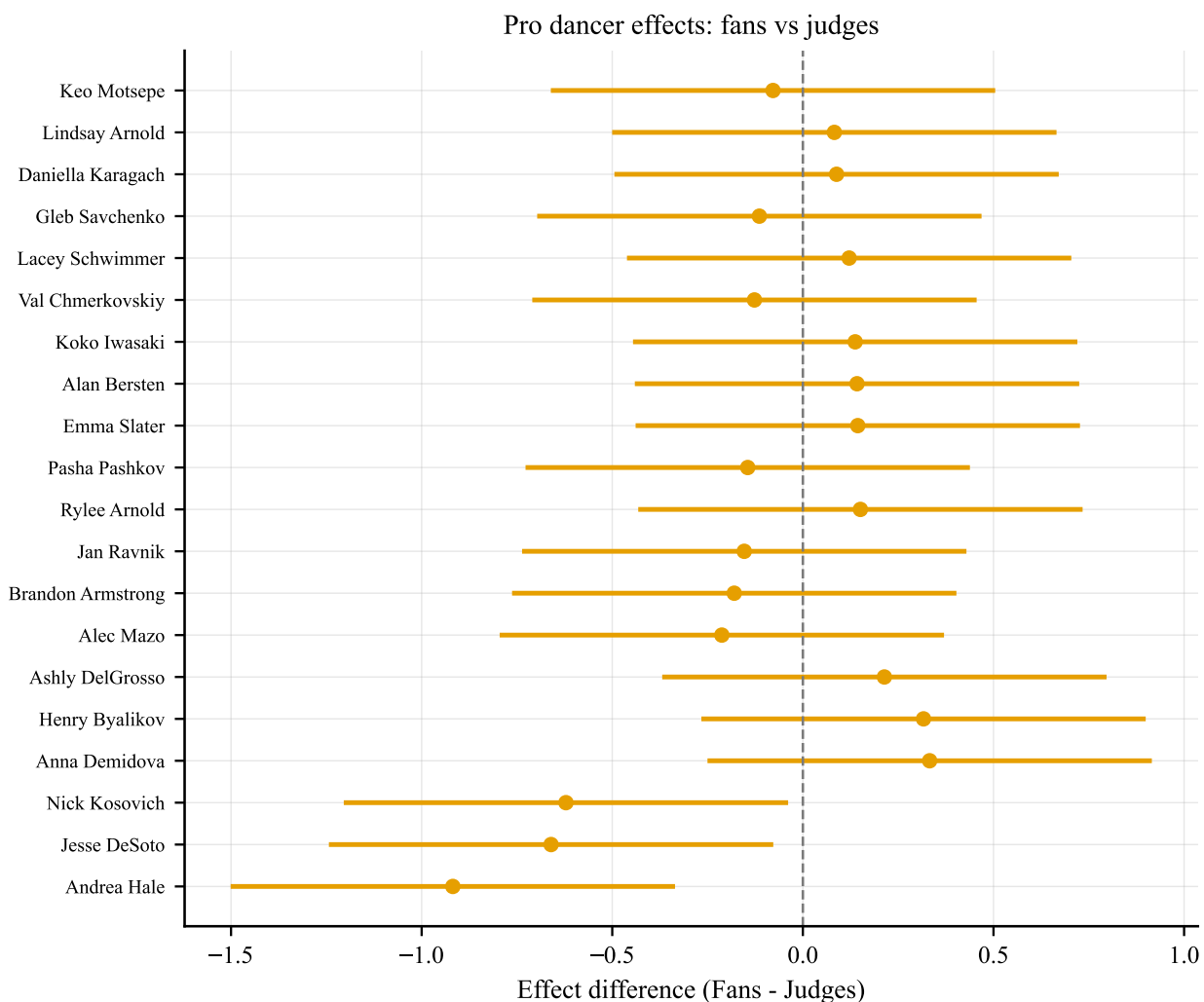


图 13: 职业舞伴效应差异（粉丝 - 评委）（正值 = 更受粉丝偏好，负值 = 更受评委偏好）。

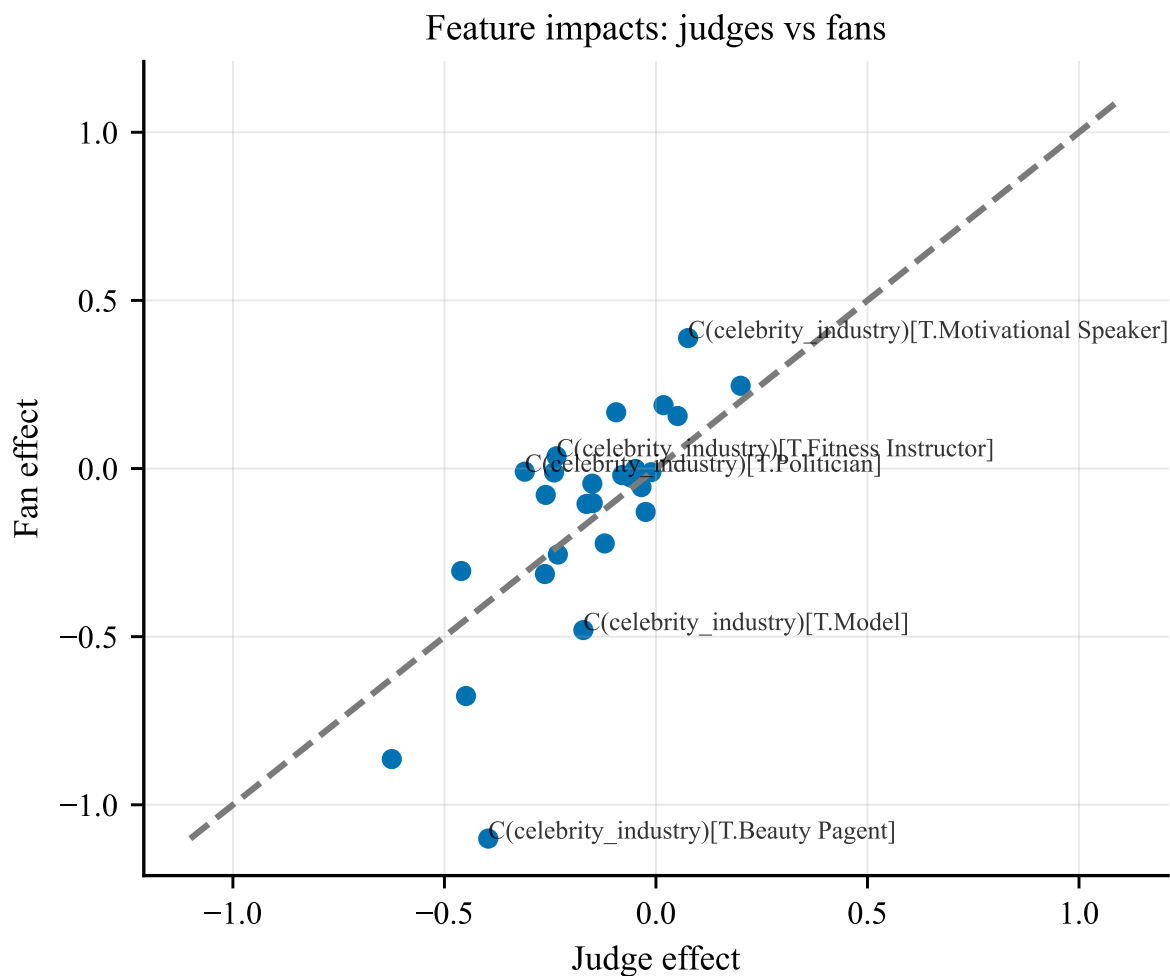


图 14: 标注离对角线最远的特征以突出差异（远离对角线 = 评委/粉丝效应不一致）。

## 7.1 预测补充：GBDT

结论要点. 预测模型仅作为协变量有效性的鲁棒性检验。

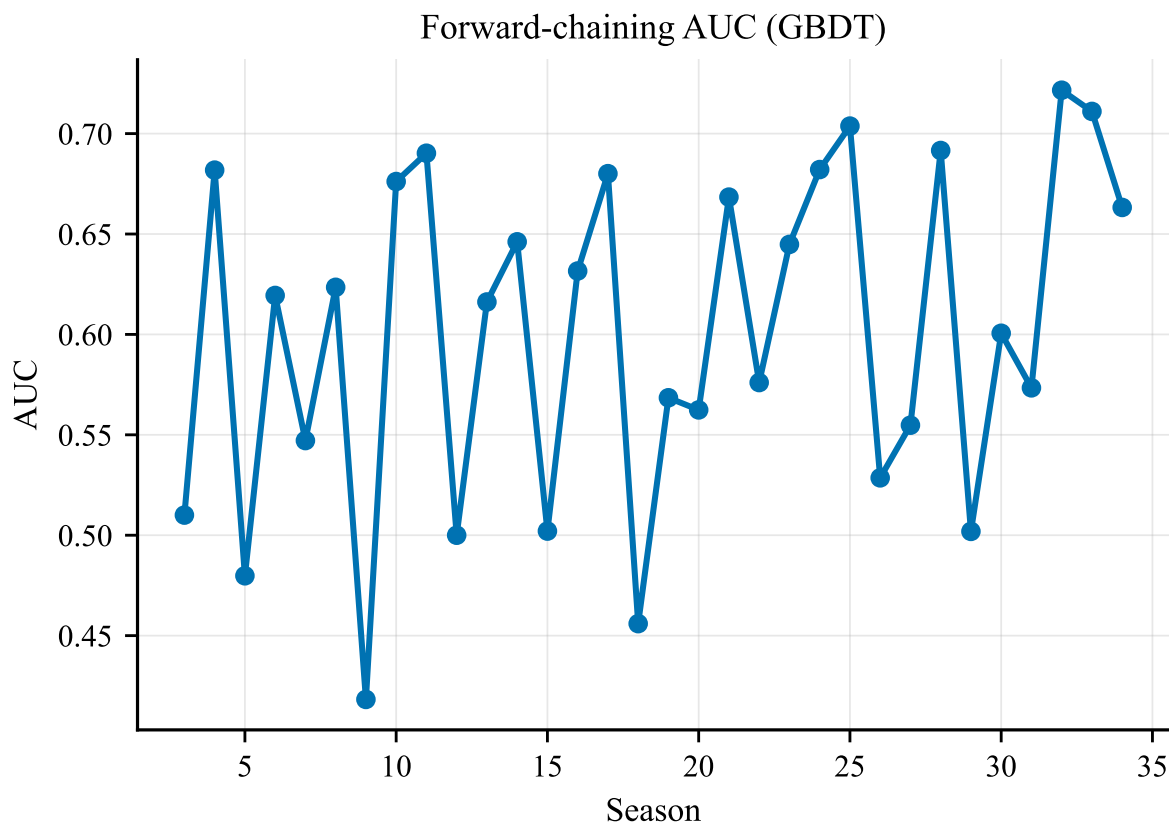


图 15: 分季 AUC 表现稳定（前向链式验证）。

关键输出. 双模型对比与任务 3 的直接回答。

## 8 模型 D: 机制设计 (DAWS)

结论要点. DAWS 以不确定性驱动权重调整, 满足单调性与稳定性。

$$\alpha_t = \text{clip}\left(\alpha_0 + \gamma \frac{t}{T} - \eta U_t, \alpha_{\min}, \alpha_{\max}\right), \quad |\alpha_t - \alpha_{t-1}| \leq \delta. \quad (13)$$

命题 1 (单调性). 评委与粉丝占比同时上升时, DAWS 得分不下降。

命题 2 (稳定性界).

$$|c_{i,t} - c_{i,t-1}| \leq \delta |j_{i,t} - v_{i,t}| + (1 - \alpha_t) \|\mathbf{v}_t - \mathbf{v}_{t-1}\| + \alpha_t \|\mathbf{j}_t - \mathbf{j}_{t-1}\|. \quad (14)$$

## 8.1 Judge-save 参数学习

$$\Pr(E = a \mid \{a, b\}) = \sigma(\beta(J_b - J_a)) \quad (15)$$

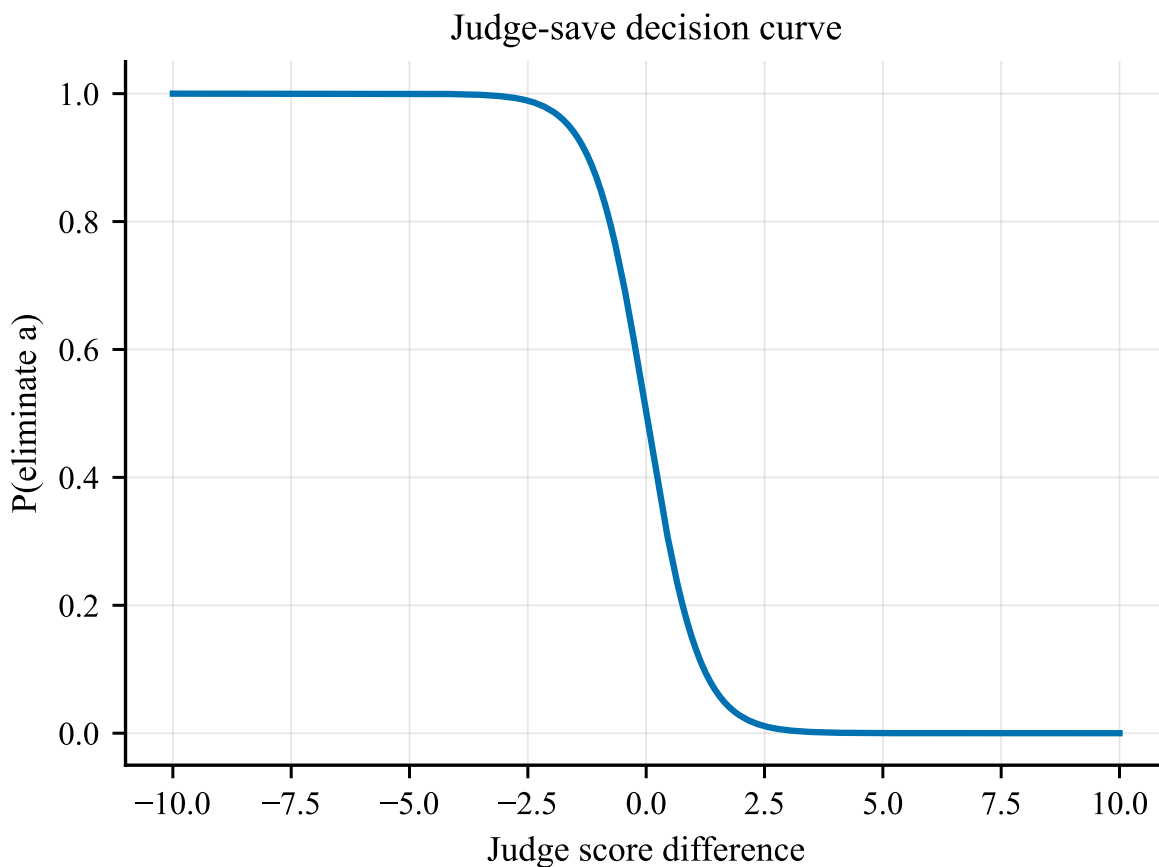


图 16: Judge-save 决策曲线（横轴评委分差，纵轴淘汰概率；示意曲线使用  $\beta = 1.8$ ，分差越大则低分选手被淘汰概率越高）。

关键输出. DAWS 方案、性质与 judge-save 行为刻画。

## 9 敏感性与验证

结论要点. 关键结论对  $\sigma$ 、 $\epsilon$  与规则切换先验具有稳健性。

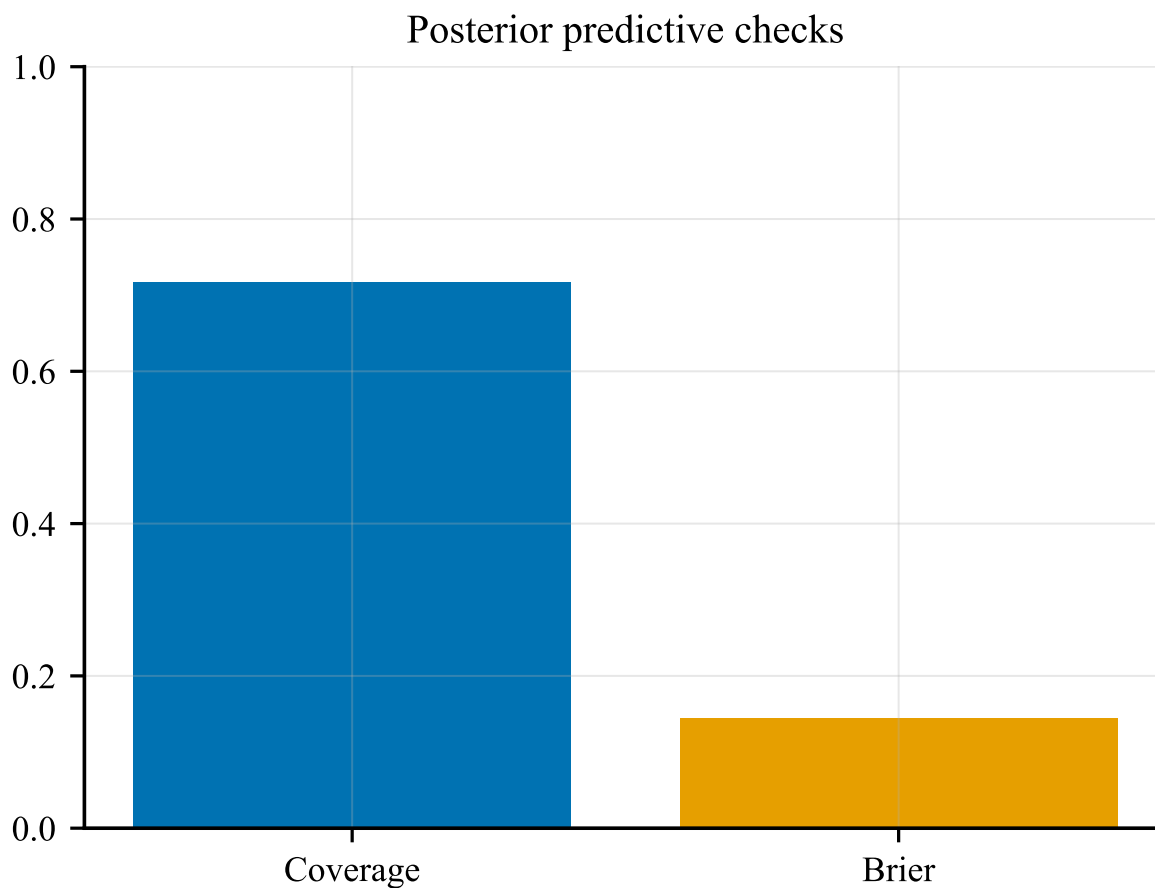


图 17: 后验预测检验结果 (Coverage 越高越好, Brier 越低越好; Brier 反映概率预测误差大小)。

## 9.1 规模对比实验

我们在多进程设置下比较不同采样规模, 记录运行时间、误差 (均值 HDI 宽度)、稳定性 (DAWS) 与理论匹配度 (Kendall  $\tau$ )。结果显示误差随规模提升而趋于平缓, 图中虚线标注了拐点与最终规模选择。



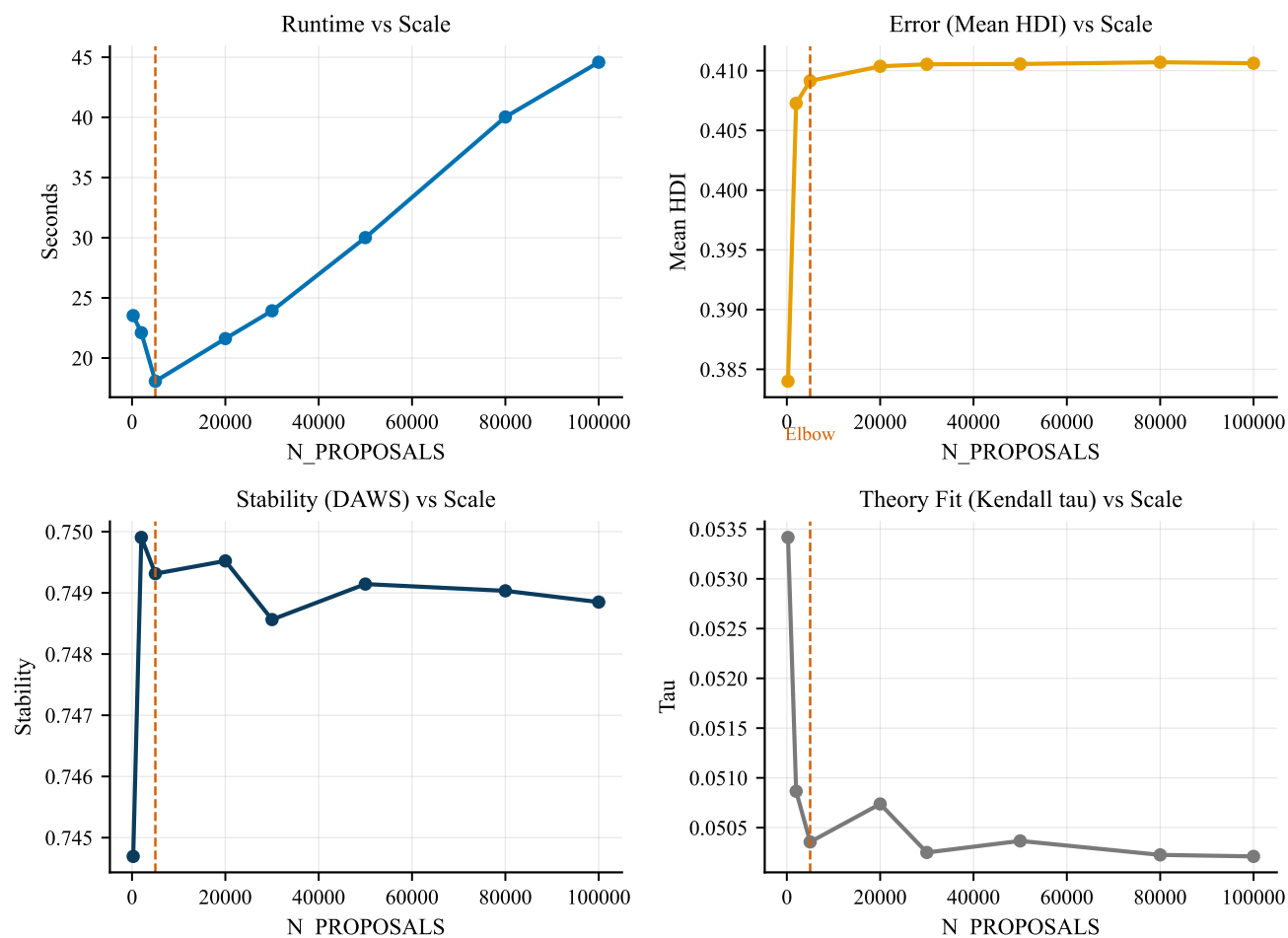


图 18: 不同采样规模下的时间、误差、稳定性与匹配度对比（虚线为折中规模）。

关键输出. 敏感性曲线与后验预测覆盖率。

## 10 结论与建议

结论要点. 审计先行揭示关键不确定性，DAWS 提供稳健折中。

我们完成全赛季粉丝票审计，量化 Rank 机制的民主赤字，并提出 DAWS 方案以提升公平、主权与稳定。

- **可读结论：**不确定性集中在少数周，其余周可识别性较高。
- **机制影响：**Rank 聚合提高翻转概率；DAWS 在稳定性上明显更优。
- **落地建议：**公布 DAWS 权重计划与 judge-save 规则以提升透明度。

## 参考文献

- [1] COMAP. 2026 MCM/ICM Problem C: Dancing with the Stars (DWTS). Contest Problem Statement.
- [2] Smith, R. (1984). Efficient Monte Carlo procedures for generating points uniformly in polytopes. *Operations Research*.
- [3] Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*.
- [4] Gelman, A., et al. (2013). *Bayesian Data Analysis*. CRC Press.
- [5] Moulin, H. (1988). *Axioms of Cooperative Decision Making*. Cambridge Univ. Press.

## AI 使用报告

我们使用 AI 协助完成论文结构草稿、LaTeX 模板与方法表述润色；所有模型选择与解释均由团队复核并最终确认。

- 可复现性：代码、图表与指标均由提供数据自动生成。
- 环境：Miniforge + mcm2026，科学计算栈已固定版本。
- 过程留痕：运行日志与汇总指标可追溯每次实验。