arXiv:2601.05914v1 [econ.TH] 9 Jan 2026

选择性披露下的贝叶斯说服*

Yifan Dai† Drew Fudenberg‡ Harry Pei§

2026年1月12日

摘要： 发送者首先公开承诺进行一项实验，之后可以私下开展额外实验，并选择性地向接收者披露实验结果。发送者掌握有关其可进行的额外实验最大数量的私人信息（即其类型）。我们证明，如果（i）接收者对发送者的类型存在足够大的不确定性，且（ii）发送者在完成完全承诺下的最优实验后仍能从选择性披露中获益，那么在任何均衡中，发送者都无法获得其承诺收益。否则，可能存在发送者获得其承诺收益的均衡。

关键词： 贝叶斯劝说、缺乏承诺、选择性披露

JEL分类代码： C73, D82, D83。

*我们感谢 Pak Hung Au、Jeffrey Ely、Ying Gao、Alexander Jakobsen、Navin Kartik、Xiao Lin、Elliot Lipnowski、Stephen Morris、Paula Onuchic、Alessandro Pavan、Refine.ink、Jean Tirole、Alexander Wolitzky、Kun Zhang 和 Gabriel Ziegler所提供的宝贵交流，并感谢美国国家科学基金会通过 SES-2337566 和 SES-2417162 项目提供的资金支持。

†麻省理工学院经济学系。电子邮箱：yfdai@mit.edu
‡麻省理工学院经济学系。电子邮箱：drew.fudenberg@gmail.com§西北大学经济学系。电子邮箱：harrydp@northwestern.edu

1 引言

专家的可信性取决于其能否对获取和披露信息的方式作出承诺。例如，一名主要寻求定罪但缺乏承诺能力的检察官，无法向法官可信地传达有用信息。但当检察官能够承诺采用某种信息收集流程，并完整披露其结果时，如 Kamenica and Gentzkow (2011) 的模型所示，其陈述便可能具有信息价值，并可能影响法官的决定。

在实践中，即使专家能够承诺获取并如实披露某些类型的信息，也可能无法承诺不开展额外调查或不隐瞒不利发现。制药行业就存在这一问题：一些临床试验会在美国食品药品监督管理局预先注册并必须报告，但公司可以在获批后开展额外试验。由于这些获批后试验未经预先注册，公司有时会隐瞒或延迟公布试验结果，尽管法律要求其报告所有试验的结果。[1] 这可能会损害专家的可信性，因为决策者（例如医生和患者）可能怀疑不利证据遭到了隐瞒，进而对专家的陈述打折扣。

受这一问题启发，我们研究一个模型：发送者（专家）公开承诺进行一项初始实验，并向接收者披露其结果。不同于Kamenica and Gentzkow （2011）的完全承诺基准，我们模型中的发送者无法承诺不私下开展额外实验并隐瞒不利结果。发送者掌握有关自己最多能够开展多少项额外实验的私人信息，这被称为其类型或能力。

我们要探讨的问题是：发送者的有限承诺何时会使其无法获得完全承诺下的收益（即其承诺收益）。我们的结果表明，答案取决于两个因素：（i）接收者对发送者的先验不确定性

[1] DeVito、Bacon 和 Goldacre(2020) 报告称，在 2018 至 2019 年间必须向美国食品药品监督管理局提交结果的 4,209 项试验中，只有 722 项（40.9%）在规定的一年期限内完成了提交。止痛药 Vioxx 是延迟报告的一个著名案例：其制造商默克推迟发布获批后试验的数据，而这些数据将该药与心脏问题联系起来。另一个案例是葛兰素史克的抗抑郁药帕罗西汀——法律文件揭露，该公司隐瞒了临床试验的证据，这些证据表明该药可能导致青少年产生自杀念头。与隐瞒或延迟披露相比，公然伪造数据的情况较少，通常也会受到更严厉的处罚。

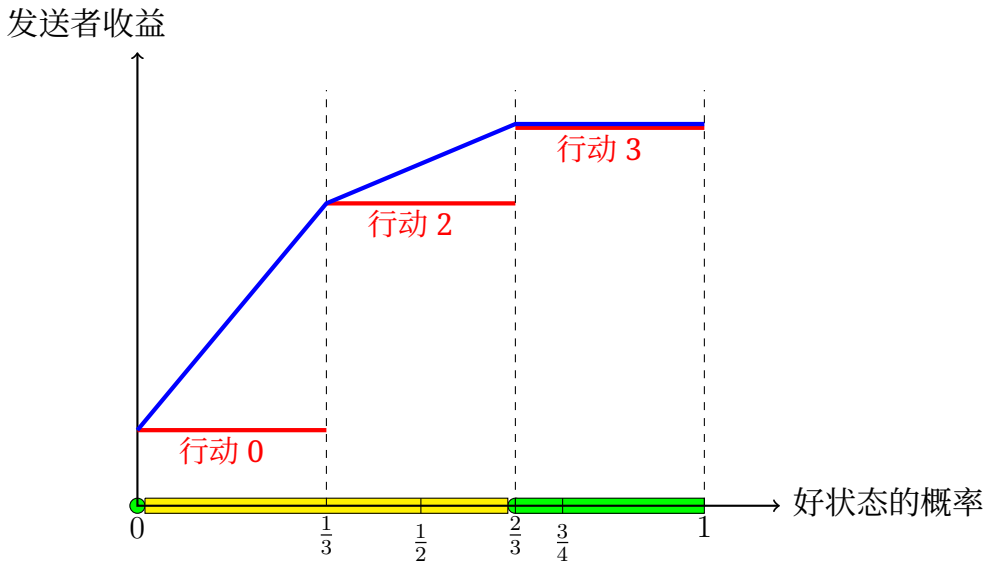

图1：发送者在接收者采取最佳回应时的效用随接收者信念变化的函数（红色）及其凹闭包（蓝色），黄色表示不可信信念，绿色表示可信信念。

类型，以及(ii)发送者在进行完全承诺下的最优实验（以下简称最优实验）后，是否能从选择性披露更多信息中获益。

为说明我们的发现，考虑图1所示的示例：有两种状态，好 和 坏，以及三个行动：0、2 和 3。发送者的效用等于接收者的行动。当好状态的概率低于 $_{1\ 3}$ 时，接收者的最优行动为 0；当好状态的概率高于 $_{2\ 3}$ 时，接收者的最优行动为 3；其他情况下则为 2。

我们首先表明，只有当接收者无法确定发送者的能力时，发送者的有限承诺才会产生影响。这是因为，当接收者知道发送者至多可以进行 t 额外实验时，发送者可以通过进行 t 无信息的额外实验并披露所有结果，证明自己没有隐瞒任何信息。

当接收者不知道发送者的能力时，发送者能否获得其承诺收益，取决于最优实验所诱导的信念，尤其取决于发送者是否能在这些信念中的某些信念下，通过选择性披露更多信息获益。在这个示例中，当某个信念赋予$_{1\ 2}$ 的概率给好状态时，发送者可以通过披露更多信息获益：他们可以进行一项额外实验

该实验会揭示状态，而发送者只披露能够证明状态为好的结果。这样一来，状态为好时，接收者的行动会从2提高到3；状态为坏时，行动则维持在2。相比之下，当某个信念赋予好状态的概率为0或$\frac{3}{4}$时，发送者无法通过披露更多信息获益，因为在前一种情况下，接收者的行动始终为0；在后一种情况下，行动永远不会超过3。我们将发送者可以通过揭示更多信息获益的信念称为不可信信念。在这个示例中，这些信念赋予好状态的概率严格介于0和$\frac{2}{3}$之间。

定理 1 表明，在如我们示例所示的单调环境 中，当且仅当最优实验不会诱导出不可信信念时，发送者才能在所有类型分布下实现其承诺收益。如果最优实验诱导出至少一个不可信信念，那么存在一组开集类型分布，使得在这些分布下，发送者的所有均衡收益都统一低于其承诺收益。

定理 2 给出了一个互补结果，且不要求单调性。该定理关注这样的环境：当状态为 θ 时，发送者严格偏好完全披露该状态，而非最优实验在状态为 θ时诱导出的任何信念。与最优实验诱导出不可信信念相比，这一条件要求更强；在好状态发生概率低于 $1/3$时，我们的示例满足后一条件。我们证明，如果有足够多的发送者类型以严格大于 0 的概率出现，那么每个均衡中的发送者收益都统一低于其承诺收益。

为了说明发送者无法实现其承诺收益，我们考察以下偏离，并据此给出较高类型均衡收益的下界：先采用较低类型的均衡策略，然后进行一次完全信息附加实验，并且仅当披露实验结果能够提高发送者收益时才予以披露。如果最优实验诱导出的信念满足我们的条件，那么在发送者实现其承诺收益的每个均衡中，较高类型通过这一偏离获得的收益都会比其较低类型的均衡收益高出一个常数。由于收益有界，一旦以正概率出现的类型足够多，就会产生矛盾。

我们还通过一个示例说明，当接收者的信念仅对少数几种发送者类型赋予正概率时，即使最优实验会诱导出某些不可信信念，发送者仍可能获得其承诺收益。接着，我们在定理 3 中证明，当

某一种发送者类型出现的概率接近 1 时，存在一种均衡，使发送者收益接近其承诺收益。

本文通过考察一种新颖的有限承诺形式，为贝叶斯劝说文献作出了贡献。现有文献大多关注发送者无法承诺进行公开实验，而我们研究的是发送者无法承诺不开展额外的私人实验。这产生了一种新的经济摩擦：接收者对未披露的后续调查的怀疑，削弱了发送者的可信性。

因此，我们的工作不同于以下模型：发送者可以承诺进行公开实验，但无法承诺如实披露其所获知的信息（Lipnowski、Ravid和Shishkin（2022））、Kreutzkamp和Lou（2025）、Lyu和Suen（2025））；或者无法承诺进行公开实验，但可以承诺不伪造信息（Henry和Ottaviani（2019）、Titova和Zhang（2025））；或者无法承诺进行公开实验，但可以通过具有承诺能力的调解人进行沟通（Myerson（1982）、Forges（1986）、Goltsman、H¨orner、Pavlov和Squintani（2009）、Corrao和Dai（2025））；或者无法承诺不偏离至诱导相同消息分布的策略（Lin和Liu（2024））；或者无法在阶段博弈中作出承诺，但会与接收者序列进行重复互动（例如，Pei（2023）、Best和Quigley（2024）、Mathevet、Pearce和Stacchetti（2024））。[2]

在Felgenhauer andSchulte(2014）和Felgenhauer and Loerke (2017）中，发送者无法承诺采用公开实验，但可以私下进行一系列二元实验，然后从所得结果中选择一个子集披露给接收者。[3] 与本文关系更密切的一篇论文，是Arieli and Stewart (2025）近期独立完成的研究。与我们的模型一样，接收者并不确定发送者可以进行多少次额外实验。该论文考察了更一般的发送者收益函数，并给出了至少存在一个弱完美贝叶斯均衡（wPBE）可实现承诺收益的条件；弱完美贝叶斯均衡是一种弱于完美贝叶斯均衡（PBE）的概念。相比之下，我们采用完美贝叶斯均衡的一种扩展；在有限信号博弈中，该扩展会化简为完美贝叶斯均衡。我们将在第4节进一步讨论不同均衡概念的影响。

[2]Lipnowski 和 Ravid (2020) 比较了发送者的承诺收益与其在发送者观察到状态的廉价交谈博弈中的收益。[3]Lou (2023) 与 Felgenhauer 和 Loerke (2017) 的研究类似，不同之处在于，一旦发送者披露某项实验的结果，就必须披露此前进行的所有实验的结果。

在我们的模型中，发送者在进行额外实验之前，会收到关于其类型的私人信息。这与有关信息设计中知情发送者的文献相关，其中包括Perez-Richet (2014）、Koessler and Skreta (2023）以及Zapechelnyuk (2023）的研究。在这些论文中，发送者的私人信息会直接影响接收者的收益，而在我们的模型中，这些信息与接收者的收益无关。[4]

我们的论文也与以下研究文献相关：该文献沿袭 Dye (1985)，研究当接收者不知道发送者可获得的证据数量时，证据披露的问题；尤其与 Dziuda(2011) 以及Gao (2025) 的研究相关。在这些研究中，发送者拥有多份证据，并决定披露其中哪些子集。在我们的论文中，发送者可获得的证据由其选择的额外实验决定。 Pei (2025) 研究了一个采用匿名随机匹配的重复博弈，其中玩家可以选择性披露由其过去行动产生的信号，且玩家不知道其搭档产生了多少信号。这一特点与我们的模型类似：只有当接收者对发送者的能力感到不确定时，发送者的有限承诺才会降低其收益。

2 基准模型

发送者（s）与接收者（r）互动。发送者的收益 $u^s(a)$ 仅取决于接收者的行动 $a \in A$，因此他们具有透明动机，这是 Lipnowski 和 Ravid(2020) 所指的意义。接收者的收益 $u^r(\theta, a)$ 同时取决于 a 和世界状态 $\theta \in \Theta$。起初，双方都不知道该状态；该状态根据一个具有全支持的分布$\pi_0 \in \Delta(\Theta)$ 抽取。我们假设 Θ 和 A 都是有限集。

一个实验 σ 是从 Θ 到结果上的概率分布集合的映射$s \in S$，其中 S 是一个可数无限集。令 Σ 表示所有实验的集合；对于每个实验 $\sigma \in \Sigma$，令 $\sigma(s \mid \theta)$ 表示在 θ条件下 s 的概率。

博弈开始时，发送者选择一个初始实验 σ_0，随后双方观察到 σ_0 及其实现的结果 $s_0 \in S$。观察到（σ_0， s_0）之后、接收者选择行动之前，发送者可以进行一些额外实验。

[4]Rappoport 和 Somma (2017)研究了一种模型：发送者拥有关于所进行的单个实验的私人信息，但这最终并不会影响接收者对状态的信念。

假设在给定 θ的条件下，不同实验的结果相互独立。发送者可进行的额外实验的最大数量为 $t \in \{0,1,...\} \equiv \mathbb{N}$，其依据 $p \in \Delta(\mathbb{N})$ 抽取。发送者在选择 σ_0 后，私下观察到 t [5]我们将 t 称为发送者的（中间）类型。

发送者在观察到（σ_0、 s_0、 t）后，依次进行额外实验，因此其对第 k 个额外实验的选择可以取决于（σ_0、 s_0、 t）、前 $k-1$ 个额外实验及其结果 $\{\sigma_i, s_i\}_{i=1}^{k-1}$。随后，发送者选择披露额外实验结果的一个子集，然后接收者选择 $a \in A$。

我们假设，当且仅当发送者披露某个额外实验的结果时，接收者才会观察到该实验及其结果。因此，接收者无法直接观察 t，即发送者进行的额外实验数量，以及未披露的额外实验的内容和结果。

令 $\mathcal{H}_k$ 表示由 k 个实验及其各自某个结果组成的序列集合。令 $\mathcal{H} \equiv \bigcup_{k=1}^{+\infty} \mathcal{H}_k$，其典型元素记为 $h \in \mathcal{H}$。接收者的信息集是 $\mathcal{H}$ 中与已披露实验相对应的元素。[6]接收者的策略 $\alpha : \mathcal{H} \to \Delta(A)$ 将其信息集映射到其行动。

发送者除初始节点外的信息集属于 $\mathcal{H} \times \mathbb{N}$。其纯策略由以下部分构成：**(i)** 一个初始实验 $\sigma_0 \in \Sigma$；**(ii)** 一个映射，将每种类型 $t \geq 1$以及每个满足 $1 \leq k \leq t$ 的 $h \in \mathcal{H}_k$ 映射到 $\Sigma \cup \{\emptyset\}$，以表示发送者根据其类型和先前实验的结果选择下一项实验的决策，其中 $\emptyset$ 表示不再进行任何实验；以及 **(iii)** 一个映射，将每种类型 $t \geq 1$ 以及每个满足 $1 \leq k \leq t+1$ 的$h \in \mathcal{H}_k$ 映射到 $\{1,2,...,k-1\}$的子集；当披露第 i 个额外实验时， i 属于所选子集。发送者的混合策略是其纯策略的一个分布。

接收者在 $h \in \mathcal{H}$ 时的朴素信念记为 $\widehat{\pi}_h \in \Delta(\Theta)$，其定义为：如果接收者观察到 h，并相信发送者已披露其进行的所有实验的结果，那么接收者对 θ 的信念。根据定义，接收者的朴素信念仅取决于 h 中的实验—结果对，而不取决于这些实验—结果对的顺序以及发送者的策略。

[5]第 4 节讨论发送者在选择初始实验之前观察到 t 的情形。[6]接收者是否观察到所披露实验的顺序，不会影响我们的任何结果。[7]类型为 t 的发送者只能选择 $\emptyset$（即不再进行任何实验），对于任何 $h \in \mathcal{H}_{t+1}$均如此。

接收者的朴素信念不一定与其后验信念一致，因为接收者可能会怀疑发送者隐瞒了某些额外实验的结果。

我们的博弈既不是有限博弈，也不是具有可观察行动的多阶段博弈，因此序贯均衡（Kreps and Wilson (1982)）或完美贝叶斯均衡（Fudenbergand Tirole (1991)）都不能直接适用。我们将改用 Fudenberg andTirole (1991)）提出的 完美扩展贝叶斯均衡 （PEBE，有时也简称均衡）作为我们的解概念。

形式上，我们博弈的一个 PEBE 包含每位玩家的一项策略，以及一个从 $\mathcal{H}$ 到 $\Delta(\Theta \times \mathbb{N})$ 的映射，用以刻画接收者对 $(\theta,\ t)$ 的信念；该映射满足：（i）在每个信息集上，给定各自的信念，每位玩家的策略都是对对手策略的最优回应；（ii）接收者的信念在均衡路径上由贝叶斯法则推导得出，并且在每个信息集上都与发送者策略所诱导的某个条件概率系统相一致。[8]

3 分析

我们考察发送者的有限承诺何时会妨碍其获得承诺收益，即其在Kamenica and Gentzkow (2011）所设基准中的收益。我们的结果表明，答案取决于两个因素：（i）接收者对发送者类型是否面临足够的不确定性；（ii）发送者能否在实施其最优实验后，通过选择性披露更多信息而获益。

3.1 假设与预备知识

现在我们引入假设1 和2，这些假设将在全文中始终成立。令 $A^*(\pi) \equiv$最大值点集 $_{a\in A} \sum_{\theta\in\Theta} \pi(\theta)u^r(\theta,\ a)$ 表示在以下条件下接收者最优行动的集合

[8]集合 Ω 上的条件概率系统（CPS）是一个函数 $\mu(\cdot|\cdot): 2^\Omega \times 2^\Omega\backslash\{\emptyset\} \to [0,1]$，满足：(i) 对所有非空 $\Omega_0 \subseteq \Omega$，$\mu(\cdot|\Omega_0)$ 是 Ω_0 上的概率分布；(ii) 对所有 $\Omega_3 \subseteq \Omega_2 \subseteq\Omega_1 \subseteq\Omega$ 且 $\Omega_2 \neq \emptyset$的情形，都有 $\mu(\Omega_3|\Omega_2)\mu(\Omega_2|\Omega_1) = \mu(\Omega_3|\Omega_1)$。在我们的博弈中，$\Omega$ 是可行终端节点的集合，因此每个条件概率系统都会将发送者进行的实验次数超过其类型允许次数这一事件赋予条件概率 0。附录A对完美贝叶斯均衡的定义作了非常小的扩展，以处理发送者选择实验会导致可验证信号形成概率分布这一事实。附录还表明，完美贝叶斯均衡具有一个直观上颇为吸引人的性质：当 t 有上界 t^*时，接收者在观察到任意 $h \in \mathcal{H}_{t^*+1}$ 后的条件信念等于其朴素信念。Arieli and Stewart（2025）所采用的解概念——弱完美贝叶斯均衡——并不要求这一性质。

$\pi \in \Delta(\Theta)$。

假设1。 对于每个 $a \in A$ 和 $\pi \in \Delta(\Theta)$，只要满足 $a \in A^*(\pi)$，就存在$\pi' \in \Delta(\Theta)$，使得 $supp(\pi') \subseteq supp(\pi)$，且满足 $A^*(\pi') = \{a\}$。

假设1要求，对于接收者在关于状态的某个信念 a 下的每个最优行动 π，都存在 π'，其支撑集弱较小，且在该信念下 a 对接收者而言是严格最优的。这意味着接收者在知晓状态时有严格最优回应。由于 Θ 和 A 均为有限集，根据假设1，对于 $u^r(\theta, a)$ 的一个满勒贝格测度集合，在 $\mathbb{R}^{|\Theta|\times|A|}$ 中成立。此结论由Lipnowski、Ravid 和 Shishkin（2025）的命题2得出。

令 $u(\theta)$表示接收者在状态 θ下采取其（唯一）最佳回应时发送者的收益。发送者的完全披露收益为 $V(\pi_0) \equiv \sum_{\theta\in\Theta} \pi_0(\theta)u(\theta)$。

如Kamenica and Gentzkow（2011）所示，一旦接收者的先验 $\pi_0 \in \Delta(\Theta)$ 固定，每个实验 σ 都由其诱导的朴素信念分布 $\pi \in \Delta(\Theta)$ 刻画，下文简称为信念。我们用 $\sigma[\pi]$ 表示信念 π 在刻画 σ的信念分布中出现的概率，并称实验σ 诱导出信念 π，若 $\sigma[\pi] > 0$。我们用 $\Sigma[\pi] \subseteq \Delta(\Delta(\Theta))$ 表示其均值等于 π的信念分布集合。令 $u(\pi) \equiv \max_{a\in A^*(\pi)} u^s(a)$ 表示接收者的信念为 π时发送者可获得的最高收益。发送者的承诺收益定义为

$$\overline{V}(\pi_0) \equiv \max_{\sigma\in\Sigma[\pi_0]} \sum_{\pi\in\Delta(\Theta)} \sigma[\pi]\overline{u}(\pi). \tag{1}$$

根据定义，$V(\underline{\pi}_0) \leq V(\overline{\pi}_0)$。若一个实验在（1）中取得最大值，则称其为最优实验。假设1 保证每个最优实验都是一个实验序列的极限，而该序列中的实验始终诱导严格的接收者激励（见附录B中的引理3）。标准的上半连续性论证表明，当发送者能够承诺不进行额外实验时，无论接收者如何打破平局，发送者都可以确保获得任意接近于 $V(\pi_0)$的收益。[9] 为简化分析，我们聚焦于（π_0, u^s, u^r）满足唯一最优实验存在的情形：

Assu假设 2. 存在唯一最优实验，记为 σ^*。

[9]因此 e，假设1导出Ali、Kleiner和Zhang（2024）中的“连续性性质”。

当 $|\Theta| = 2$时，假设2 对所有（π_0， u^s， u^r）均成立，但以下情形除外：V（π_0） $=$ 最大值$_{\pi\in\Delta(\Theta)}$ u（π）–，以及一组零测度的（u^s， u^r） $\in \mathbb{R}^{|\Theta|\times|A|^2}$。

命题 1 表明，除非 (i) 发送者的承诺收益高于其完全披露收益，且 (ii) 接收者对发送者进行额外实验的能力存在非平凡的不确定性，否则发送者在所有均衡中都能获得其承诺收益。

命题 1。

1. 对于任意 $p \in \Delta(\mathbb{N})$，发送者在每个完美贝叶斯均衡中的收益都属于 $[V(\pi_0)\ V(\pi_0)]$。

2. 对于任意退化的 $p \in \Delta(\mathbb{N})$，发送者在每个完美贝叶斯均衡中的收益都是 (π_0)。

证明见附录 C。直观而言，发送者可以在初始实验中完全披露状态，从而确保获得 $V(\pi_0)$ 的收益。当接收者知道 t 时，发送者可以通过 (i) 选择一个接近 $V(\pi_0)$ 的初始实验，使接收者在每个诱导出的信念下都有严格激励；根据假设 1，这样的实验存在；并且 (ii) 进行 t 非信息性额外实验并披露其结果，从而确保获得任意接近 σ^* 的收益。

3.2 主要结果

我们的主要结果考察了发送者的有限承诺何时会阻止其在所有均衡中获得其承诺收益。我们首先引入可信信念的定义：

定义1。 关于状态的信念 $\pi \in \Delta(\Theta)$ 若对于每个 $\theta \in supp(\pi)$，都有 $u(\pi) \geq u(\theta)$，则称其为可信。否则，该信念为不可信。

正如我们在初始示例中所说明的，退化信念以及发送者获得其最大效用时的信念都是可信的。定理1 表明，在单调的收益环境中——这类环境一直是说服文献的主要关注点——当且仅当最优实验 σ^* 不诱导出不可信信念时，发送者才能在所有类型分布下获得其承诺收益。

定义2。 (u^s, u^r) 若分别在 Θ 和 A上存在完全序，使得 $u^s(a)$ 关于 a 严格递增，且 $u^r(\theta, a)$ 具有严格递增差异，则称其为单调的。

定理1。 假设(u^s, u^r) 是单调的。

1. 若 σ^* 诱导的所有信念都是可信的，则对于每个 $p \in \Delta(\mathbb{N})$，都存在一个完美贝叶斯均衡，其中发送者收益为 $V(\pi_0)$。

2. 如果 σ^* 诱导出至少一个不可信信念，那么存在 $\eta > 0$ 和一个类型分布的开集（在积拓扑下），使得对于该集合中的每个类型分布 p，发送者在任何完美贝叶斯均衡中的收益都不超过 $V(\pi_0) - \eta$。

在引言中的示例里，当好状态的先验概率处于 $(0, 2/3)$ 时，最优实验会诱导出一个不可信信念。因此，根据定理 1，对于一个开集中的类型分布，发送者的收益严格低于承诺收益。相比之下，在Kamenica and Gentzkow (2011)的检察官—法官例子中，最优实验诱导出的所有信念都是可信的。因此，无论类型分布如何，发送者都能在某个均衡中获得承诺收益。

为证明第一个结论，假设发送者选择 σ^* 作为初始实验，并且在均衡路径上不进行额外实验；接收者在平局时倾向于发送者；并且接收者在每个信息集上的后验信念都等于其朴素信念。在单调环境中，这是一个均衡，因为 $u(\cdot)$ 在一阶随机占优序下递增：如果 $\pi \in \Delta(\Theta)$ 满足对于每个 $\theta \in \text{supp}(\pi)$，都有 $u(\pi) \geq u(\theta)$，那么对于满足 $\text{supp}(\pi') \subseteq \text{supp}(\pi)$ 的每个 π'，我们都有 $u(\pi) \geq u(\pi')$。因此，对于由 σ^*诱导的任一可信信念 π，发送者都无法通过披露任何额外信息并将接收者的后验信念转移到任意满足 $\text{supp}(\pi') \subseteq \text{supp}(\pi)$ 的 π' 而严格获益。[10]

我们对第二个结论的证明使用了一个适用于单调环境之外情形的引理：若 σ^* 诱导出至少一个不可信信念，那么对于每个足够大的 $n \in \mathbb{N}$，当 p 接近 $\{0, 1, ..., n\}$上的均匀分布时，发送者无法达到 $V(\pi_0)$。

引理 1. 若 σ^* 诱导出一个不可信信念，则存在 $\eta > 0$、$n \in \mathbb{N}$以及一个包含 $\{0, 1, ..., n\}$ 上均匀分布的类型分布开集，使得对于该开集中的每个 p，任何完美贝叶斯均衡中的发送者收益至多为 $V(\pi_0) - \eta$。

证明见附录D。证明的关键步骤是，从下方对类型 $t+1$ 发送者收益进行界定

[10]利用同样的构造性证明，可以证明，在非单调环境中，若最优实验σ^* 所诱导的每个 $\pi \in \Delta(\Theta)$ 都满足更强的可信性概念，则对于每个 $p \in \Delta(\mathbb{N})$，都存在一个完美贝叶斯均衡，使发送者收益为 $V(\pi_0)$：对于每个满足 $\text{supp}(\pi') \subseteq \text{supp}(\pi)$ 的 π'，均有$u(\pi) \equiv \min_{a \in A^*(\pi)} u^s(a)$。在单调环境中，这一更强的概念等同于可信性。

具体做法是计算以下情形下的收益：(i) 一开始采用类型 t 的均衡策略；(ii) 对类型 t所诱导的每个不可信后验信念 π ，进行一项能够完全揭示状态的额外实验；(iii) 当且仅当相应状态 θ 满足 $u(\theta) > u(\pi)$时，披露实验所得结果。由此可得类型 $t+1$与类型 t的均衡收益之差的下界，该下界是一个关于类型 t的均衡策略诱导不可信信念的概率的线性函数。反设类型数量很多，且存在一个发送者事前期望收益接近 $V(\pi_0)$的均衡。那么，鉴于最优实验的唯一性，会有许多类型以严格大于零的概率诱导不可信后验信念。由于可行收益集合有界，当以正概率出现的类型数量足够多时，就会导致矛盾。

引理 1 的证明依赖于接收者的信念所具有的一种离散性质， p 即接收者对发送者类型的信念会以有上界的正概率赋予足够多的类型。事实上，在许多自然情形下，对发送者类型的任何足够离散的信念，都足以阻止发送者获得其承诺收益。对于每个 $\theta \in \Theta$，令 $\Pi_\theta^* \subseteq \Delta(\Theta)$ 表示由 σ^* 诱导、且将 θ 纳入其支撑集的信念集合。

定理 2. 假设存在一个状态 $\theta \in \Theta$ ，使得 $u(\theta) >$ 大于$\pi \in \Pi_\theta^* u(\pi)$的最大值。那么存在一个常数 $\lambda > 0$ ，使得对于所有足够小的 $\eta > 0$，都存在 $n \in \mathbb{N}$，从而对于任意 $p \in \Delta(\mathbb{N})$，只要其赋予至少 n 个不同类型的概率均大于 $\lambda\eta$ ，发送者在每个完美贝叶斯均衡中的收益就不超过 $V(\pi_0) - \eta$。

定理2 要求存在一个状态 θ 使得，只要接收者的信念属于 Π_θ^*，发送者就严格偏好于完全揭示 θ，即 $u(\pi) < u(\theta)$。在引言中的示例里，当好状态发生的概率介于 0 和 $1\ 3$之间时，这一附加要求便得到满足。这一要求比引理 1 中的假设更强；该假设只要求 σ^* 至少诱导出一个不可信信念，因为 Π_θ^* 中的所有信念都是不可信的。附录 F 给出了一个示例，说明这一更强要求不能弱化为引理 1中的要求。

定理2的证明见附录E，其论证与引理1类似：我们再次考察高类型发送者模仿低类型的均衡策略所能获得的收益，以及

随后进行一项能够完全揭示状态的额外实验，并且仅当实验结果带来严格更高的收益时才披露该结果。当 $u(\theta) > \max_{\pi \in \Pi_\theta^*} u(\pi)$ 时，无论低类型的策略导致不可信信念的概率如何，高类型都能保证获得严格高于任何低类型均衡收益的收益。其逆命题——即存在一个均衡，使发送者收益任意接近承诺收益——在足够多种类型以正概率出现时必定不成立，由此完成论证。

接下来，我们解释为什么接收者对发送者类型持有分散的信念并非多余。我们首先给出一个示例：即使（i）最优实验诱导出不可信信念，且（ii）接收者对发送者类型面临非平凡的不确定性，发送者仍能在某个均衡中获得其承诺收益。随后，我们证明定理 3，该定理表明，当 p 集中于单一类型时，总存在一个均衡，使发送者收益接近 $V(\pi_0)$。

示例： 假设 $A \equiv \{0, 2, 3, \frac{7}{2}\}$, $\Theta \equiv \{\theta_0, \theta_1\}$ 和 $t \in \{0, 1\}$。我们将 θ_1 的概率用来表示接收者对状态的信念。发送者收益等于 a。接收者的收益满足：当 $\pi \in [0, \frac{1}{4}]$ 时，$a = 0$ 是最优的；当 $\pi \in [\frac{1}{4}, \frac{1}{2}]$ 时，$a = 2$ 是最优的；当 $\pi \in [\frac{1}{2}, \frac{3}{4}]$ 时，$a = 3$ 是最优的；当 $\pi \in [\frac{3}{4}, 1]$ 时，$a = \frac{7}{2}$ 是最优的。假设 $\pi_0 = \frac{3}{8}$，并且 $t = 0$ 的概率为 $\frac{1}{3}$。[11] 我们构造一个均衡，使发送者的期望收益为 $V(\pi_0) = \frac{5}{2}$；该收益通过以相等概率诱导后验信念 $\frac{1}{4}$ 和 $\frac{1}{2}$ 而实现。

在这个均衡中，发送者选择无信息初始实验。类型 1 进行额外实验，结果 s 在 $\theta = \theta_1$ 时以概率 1 出现，在 $\theta = \theta_0$ 时以概率 $\frac{3}{5}$ 出现，并且只披露 s。观察到 s 后，接收者的后验信念为 $\frac{1}{2}$；未观察到额外信号后，接收者的后验信念为 $\frac{1}{4}$，这与接收者在所有非均衡路径信息集上的朴素信念一致。

在此示例中，类型 1 没有动机通过额外实验完全揭示状态 θ_1。相反，他们更愿意披露一个部分揭示状态并诱导出后验信念 $\frac{1}{2}$ 的信号。当接收者只对类型 0 和类型 1 赋予正概率时，这会导向一个均衡，使发送者获得其承诺收益。尽管如此，我们的构造

[11] 更一般地，对于每个 $\pi_0 \in (\frac{1}{4}, \frac{1}{2})$，在 $t = \{0, 1\}$ 上存在一个非退化分布，并且存在相应的均衡，使发送者获得 $V(\pi_0)$。

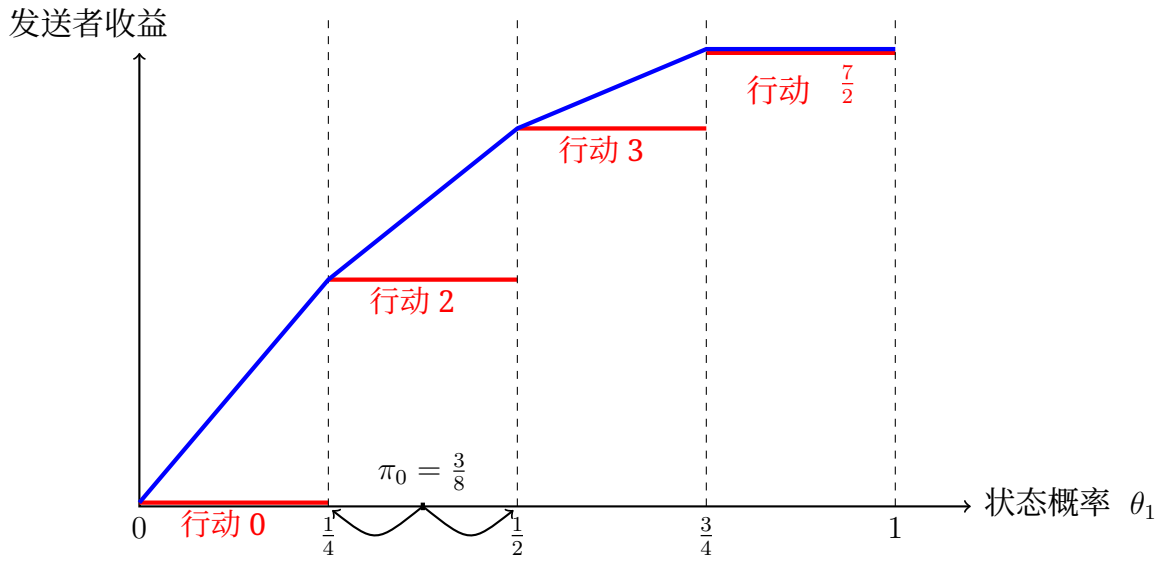

图 2：示例中发送者的间接效用随接收者信念 $u(\pi)$ 的变化（红色），以及其凹闭包（蓝色）。

依赖于仅有两种发送者类型这一假设。正如定理2所示，当发送者类型分布足够分散时，不存在这样的均衡。

接下来，我们证明，如果类型分布 p 将足够接近 1 的概率赋予某一类型，那么存在一个均衡，使发送者的收益接近 $V(\pi_0)$。

定理3。对于每个 $\delta > 0$，存在 $\varepsilon > 0$，使得对于每个 $p \in \Delta(\mathbb{N})$，其中$p(n) > 1 - \varepsilon$ 对于某个 $n \in \mathbb{N}$，存在一个完美贝叶斯均衡，使发送者的收益大于$V(\pi_0) - \delta$。

我们的三个定理共同表明，当收益单调时，只要最优实验诱导出某些不可信信念，发送者能否在其最优均衡中获得承诺收益，就取决于接收者对发送者类型的信念的离散程度。如果该信念足够集中于单一类型（因此接收者对发送者能力的不确定性很小），那么就存在一些均衡，使发送者的收益约为 $V(\pi_0)$。相反，如果接收者的信念赋予多个发送者类型正概率，且每种此类类型出现的概率都有一个严格大于 0 的上界，那么发送者的均衡收益就严格低于其承诺收益。

定理2 和定理3 互为补充，因为它们的证明分别给出了发送者的承诺收益与其最高均衡收益之差的下界和上界

。定理2 的证明表明，该差距的下界是最可能类型的概率的线性函数 n，其中 n 是以下因素的函数：发送者在不可信信念下选择性披露额外信息所获得的收益、关于状态的先验信念，以及诱导这些不可信信念的最优实验以该概率发生的概率。定理3 则给出了该差距的上界，它是除最可能类型之外的其他类型的概率之和的线性函数。

定理 3的证明见附录 G，其中构造了可近似实现发送者承诺收益的均衡。为便于理解，考虑类型 0 出现概率为零这一简单情形。在我们的构造中，发送者选择一个无信息初始实验，并且：(i) 当其实现类型介于 1 和 $n-1$之间时，进行一次揭示状态的额外实验，并完整披露结果；(ii) 当其实现类型至少为 n时，进行一次接近 σ^*的额外实验，完整披露其结果以及 $n-1$ 无信息实验的结果；(iii) 当其实现类型高于 n时，可能进行多于 n 次额外实验，并选择性披露其结果；(iv) 在每个非均衡路径信息集上，接收者的后验信念都以概率 1 将其朴素信念支撑集中的最差状态赋予最高信念，即使 θ 最小化 $u(\theta)$。由于类型 $V(\pi_0)$ 出现的概率接近 1，发送者的事前期望收益接近 n 。

—

当类型 0 以正概率出现时，这一构造无法奏效，因为在未披露任何额外结果时，接收者会将类型为 0 的概率赋为 1，这可能会促使类型 1 至 $n-1$ 不披露完全信息附加实验的结果。我们的证明调整了类型 1 至 n 的策略，以适应这一情况。在均衡中，发送者选择一个接近 σ^* 的实验作为其初始实验。每种类型所进行的额外实验以及所披露的结果，取决于类型 0 与类型 1 至 $n-1$的相对概率。当相较于类型 1 至 $n-1$而言，类型 0 出现的可能性较高时，类型 n 会以正概率模仿这些较低类型，从而降低较低类型的均衡收益。当相较于类型 1 至 $n-1$而言，类型 0 出现的可能性较低时，类型 1 至 $n-1$ 会进行一个完全揭示状态的额外实验，并且当且仅当其在该状态下的间接效用足够高时才披露结果。

4 讨论

我们首先给出一个示例，说明 PEBE 的路径外限制——而 wPBE 不具备这些限制——如何迫使发送者的均衡收益低于其承诺收益。随后，我们考察发送者的最差均衡收益，以补充我们先前对最佳均衡的分析。最后，我们考虑一些扩展情形：发送者在选择初始实验之前观察到自己的类型，或者发送者为每次额外实验承担成本。

完美贝叶斯均衡 与 弱完美贝叶斯均衡： 考虑以下示例： $A = \{0, 2, 3\}$, $\Theta = \{\theta_0, \theta_1\}$, $t \in \{0, 1\}$。接收者的先验对状态 θ_1 赋予严格介于0和1 3 之间的概率，并对类型t = 1赋予 1 2 的概率。发送者的收益等于 a。令 $\pi \in [0, 1]$ 表示接收者的后验信念赋予状态 θ_1的概率。我们假设，当$\pi \in [0, \frac{1}{3}]$时，接收者选择行动0是最优的；当 $\pi \in [\frac{1}{3}, \frac{2}{3}]$时，选择行动2是最优的；当 $\pi \in [\frac{2}{3}, 1]$时，选择行动3是最优的。

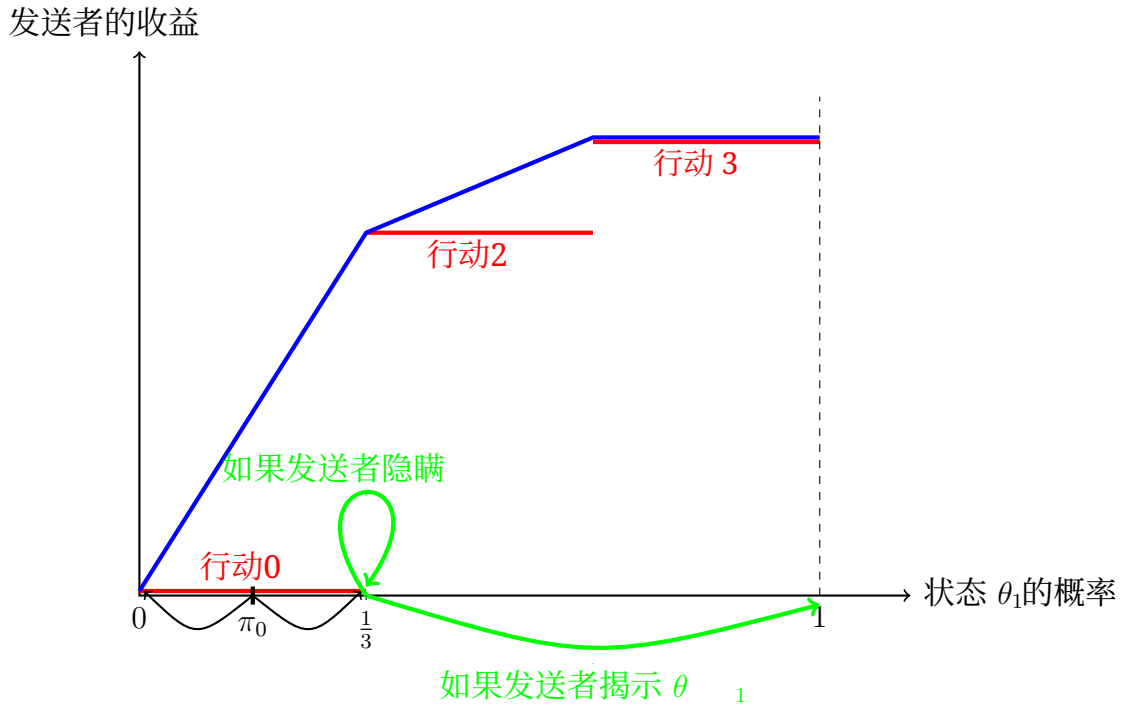

图3：发送者收益函数 $u(\pi)$ (红色)及其凹闭包(蓝色)。绿色箭头表示类型 1 的发送者在观察到初始实验的结果后，采取有利可图的偏离，从而得到朴素信念 1 3。

我们证明，在每个完美贝叶斯均衡中，发送者收益都严格低于 $V(\pi_0)$。反设

存在一个完美贝叶斯均衡，使发送者收益等于 $V(\pi_0)$。那么，对于每种发送者类型，在状态 θ_1 下，接收者的后验信念必定为1 3 。由于两种类型出现的概率都为正，因此存在一个 $h \in \mathcal{H}_1$ ，使得接收者的后验信念为 1 3。但此时，类型 1 严格有动机在 h 进行完全信息实验，并且当且仅当 $\theta = \theta_1$ 时揭示结果；这会使接收者在状态 θ_1 下的行动从 2 增至 3，而在状态 θ_0 下的行动保持不变。因此，不存在发送者获得其承诺收益 $V(\pi_0)$ 的完美贝叶斯均衡（见图3）。

—

上述论证利用了完美贝叶斯均衡与弱完美贝叶斯均衡之间的一个关键差异。在完美贝叶斯均衡下，只要发送者在任何信息集（无论是否位于均衡路径上）披露了能揭示状态 θ 的证据，接收者的后验信念就必须以概率 1 赋予 θ。这一限制在弱完美贝叶斯均衡下未必成立。[12] 正因如此，Arieli and Stewart (2025) 得以证明：对于一类收益环境（包括所有具有透明动机的情形），在某些弱完美贝叶斯均衡中，发送者能够获得其承诺收益。

发送者的最差均衡收益：我们的下一个结果表明，当发送者的类型无界时，存在一个均衡，使发送者获得其完全披露收益 $V(\pi_0)$，这是命题1中所确立的发送者均衡收益下界。

—

命题2。 若 $p \in \Delta(\mathbb{N})$ 对无穷多个类型赋予正概率，则存在一个完美贝叶斯均衡，在其中发送者完全揭示 θ ，且在均衡路径上不进行额外实验。

在接收者的信念以接近 1 的概率赋予某一发送者类型，同时也以严格为正但很小的概率赋予无穷多个类型的情形下，定理 3和命题 2 共同表明，存在一些均衡，使发送者能够近似获得其承诺收益，尽管也存在一些均衡，使发送者获得其完全披露收益。因此，相较于接收者知道发送者类型的情形（在该情形下，发送者在所有均衡中都获得其承诺收益），一旦发送者类型可能任意大，发送者的福利就取决于均衡参与者所协调的均衡，即使接收者几乎确定发送者的类型也是如此。

[12]除了刚才讨论的限制外，完美贝叶斯均衡还施加了弱完美贝叶斯均衡没有的进一步约束。例如，当发送者类型的上界为 t 时，在 t 额外实验结果被揭示的任何信息集上，接收者关于 θ 的后验信念必须等于接收者的朴素信念。

我们对命题 2 的证明构造了一个均衡，其特点类似于 Matthews and Postlewaite (1985) 中的完全揭示均衡：在状态未被完全揭示的所有信息集处，接收者形成怀疑性信念：在每个信息集$h \in \mathcal{H}$处，接收者的后验信念将以概率 1 赋予其朴素信念的支撑集中的最差状态 $\widehat{\pi}_h$（即在该状态下，接收者的唯一最佳回应使发送者收益最小化）。当发送者类型分布 p 无界时，无论发送者已经揭示了多少额外实验，接收者总会认为以下情况是合理的：发送者隐瞒了至少一个额外实验的结果，而该结果揭示了接收者朴素信念支撑集中的最差状态。

命题*2*的证明：令 $\Theta \equiv \{\theta_1, \theta_2, ..., \theta_n\}$。不失一般性，我们假设$u(\theta_n) \geq ... \geq u(\theta_1)$。考虑如下策略组合和信念：发送者选择一个完全揭示状态的初始实验，并且在$\tilde{}$ $\tilde{}$均衡路径上不进行额外实验。对于每个 $k \geq 1$ 以及每个非均衡路径历史 $h_k \in \mathcal{H}_k$，令 $i(h_k)$ 表示使得 $\theta_i \in \text{supp}(\widehat{\pi}_{\tilde{h}_k})$ 成立的最低 i。发送者类型 $t \geq k$ 将在 h_k 进行一次完全信息性的额外实验$\tilde{}$，并披露结果，除非状态为 $\theta_{i(\tilde{h}_k)}$。在每个$\tilde{}$非均衡路径信息集 $h_k \in \mathcal{H}_k$ 上，接收者的后验信念以概率 1 将状态赋为$\theta_{i(\tilde{h}_k)}$，即该信息集上接收者朴素信念支撑集中对发送者而言最差的状态。可以验证，这是一个完美贝叶斯均衡，在该均衡中，均衡路径上接收者完全获知状态，且发送者的期望收益为 $V(\pi_0)$。

□

发送者在进行初始实验之前观察其类型：在基准模型中，发送者在选择初始实验后观察其类型 t。现在我们将证明，只要关注发送者的事前期望收益，我们的所有结果都适用于发送者在选择初始实验之前观察 t 的情形。

命题1所扩展的论证分为三个步骤。首先，无论发送者何时观察到自己的类型，他们都可以通过承诺进行一个完全揭示 θ的初始实验来获得完全披露收益，从而使接收者随后产生的怀疑变得无关紧要。其次，在均衡中，发送者的策略会诱导出一个接收者信念分布，其期望为 π_0，因此在任何均衡中，发送者的事前期望收益都不可能超过$V(\pi_0)$。第三，当 p 在类型 t处为退化的 $V(\pi_0$时，发送者可以通过选择

一个接近于 σ^* 的初始实验 σ^ε （其中接收者在每个诱导出的信念下都具有严格激励），并进行 t 无信息的额外实验，同时披露所有结果。

定理2 和定理1 的第二个结论，是通过界定高类型发送者与低类型发送者的均衡收益之差得到证明的：具体而言，考察高类型发送者采取偏离策略时的收益，该偏离策略首先使用低类型发送者的均衡策略，随后进行额外实验并选择性披露由此产生的结果。我们的论证不依赖于不同类型的发送者进行相同初始实验这一假设，这意味着当发送者在选择 t 之前 σ_0观察到其类型时，我们的结论仍然成立。

定理3 和定理1 的第一个结论仍然成立，因为仍然存在一个均衡，使发送者的所有类型都选择相同的初始实验。鉴于在我们的基准模型中，每种发送者类型获得的收益至少等于其完全披露收益，如果接收者在发送者偏离并选择不同初始实验的任何信息集上具有怀疑性信念，那么它们将选择进行相同的初始实验。

命题2 也仍然成立，因为当 p 为无限多个类型赋予正概率时，仍然存在一个均衡，其中：(i) 发送者的所有类型都选择一种能够完全揭示状态的初始实验；(ii) 在每个信息集上，接收者的后验信念都以概率 1 指向其朴素信念支撑集内使发送者收益最小的状态。

额外实验的恒定边际成本：假设发送者每进行一次额外实验，都要承担一个公开已知的恒定边际成本 c 。当 c 足够小时，我们的主要结果仍然成立。具体而言，命题1 和2的结论依然成立，因为发送者可以选择完全信息的初始实验，确保获得完全披露下的收益；当 p 无界时，只要接收者持有怀疑性信念，发送者就可以在均衡中获得完全披露下的收益，因为选择完全信息的初始实验且不进行任何额外实验构成一个均衡。对于任意退化的 p 和 $\varepsilon > 0$，都存在 $c > 0$ ，使得当 $c < c$时，在所有均衡中，发送者的收益都大于 V（π_0）$-\varepsilon$ 。这是因为当 c 足够小时，发送者可以通过选择一个接近于 ε的初始实验，确保获得一个与 V（π_0）相差不超过 σ^*

此时，接收者在每个诱导出的信念下都有严格激励（根据假设1，这样的信念存在），进行 t 无信息的额外实验，并披露所有结果。

定理1 的第一项结论 $c > 0$ 适用于所有 $c > 0$，因为我们的构造性证明不要求发送者在均衡路径上进行额外实验。定理1 的第二项结论以及定理2 在 c 相对于高类型与低类型发送者之间收益差异的界限较小时仍然成立。这些界限仅取决于收益环境（π_0，u^s，u^r）以及发送者均衡收益与承诺收益之间的差距。

定理3 的类似结论 $\delta > 0$也成立：对于每个 $\delta > 0$，都存在 $\varepsilon > 0$，使得对于每个 $c < \varepsilon$ 和 $p \in \Delta(\mathbb{N})$，其中 $p(n) > 1 - \varepsilon$ 对于某个 $n \in \mathbb{N}$成立，都存在一个完美贝叶斯均衡，使发送者的收益高于 $V(\pi_0) - \delta$。—

5 结论

我们研究一种新颖的有限承诺形式：发送者承诺进行初始实验，但无法承诺不进行额外实验并选择性披露其结果。我们表明，发送者能否获得其承诺收益取决于两个关键因素：(i) 接收者关于发送者进行额外实验能力的信息的精确程度；以及 (ii) 在最优实验诱导的信念下，发送者能否从选择性披露额外信息中获益。当发送者的最优实验诱导出至少一个信念，使得发送者能够从披露更多信息中获益（即不可信信念），且接收者对发送者进行额外实验的能力具有足够大的不确定性时，在任何均衡中，发送者都无法获得接近其承诺收益的任何收益。如果环境是单调的，并且 σ^* 诱导的所有信念都是可信的，则在某个均衡中，发送者可以准确获得其承诺收益；如果接收者几乎确定发送者的类型，则存在一个均衡，使得发送者获得近似于其承诺收益的收益。

附录

具有随机后继节点的博弈中的完美贝叶斯均衡

在本附录中，我们将完美贝叶斯均衡（PEBE）的定义扩展到终止节点和后继节点具有随机性且受参与者行动影响的博弈。

考虑一个扩展式博弈，其节点为 $x \in X$，信息集为 $h \in H$，在信息集 h 上可采取行动$a \in A(h)$，信号为 $s \in S$。我们用 $h(x)$ 表示包含 x 的信息集。对于每个 $x \in X$ 和 $a \in A(h(x))$，信号 $s \in S$ 以概率 $q(s \mid x,\ a)$ 实现，从而到达后继节点 $(x,\ a,\ s) \in X$。令 Z 表示终端节点集合。给定 $Y \subseteq X$，我们用 $Z(Y)$ 表示由 Y 中某个元素作为前驱节点的终端节点子集。

定义在 Z 上的条件概率系统（CPS）是一个函数 $\nu(\cdot \mid \cdot) :\ 2^Z \times 2^Z \setminus \{\emptyset\} \to [0,1]$，满足：(i) 对所有非空 $\Omega_0 \subseteq Z$，$\nu(\cdot \mid \Omega_0)$ 是 Ω_0 上的概率分布；(ii) 对所有 $\Omega_3 \subseteq \Omega_2 \subseteq \Omega_1 \subseteq \Omega$ 且 $\Omega_2 \neq \emptyset$，均有 $\nu(\Omega_3 \mid \Omega_2)\,\nu(\Omega_2 \mid \Omega_1) = \nu(\Omega_3 \mid \Omega_1)$。定义在 Z 上的 CPSμ 诱导出条件信念 $\mu(\cdot \mid \cdot) :\ \mu(A \mid B) := \nu(Z(A) \mid Z(B))$，其中每个 $A,\ B \subseteq X$ 均满足 $Z(A) \subseteq Z(B)$。

给定策略组合 π，令 $\pi(a \mid h)$ 表示在信息集 h 采取行动$a \in A(h)$ 的诱导概率。若策略组合 π 和定义在终端节点上的 CPS ν 构成一个扩展评估，则称之为扩展评估。

定义3. 令 μ 表示与 ν相关联的条件信念。扩展评估(ν, π) 是一个完美贝叶斯均衡，当且仅当

1. π 是对条件信念 μ的最佳响应。

2. 对于每个信息集 $h \in H$、行动 $a \in A\,(h)$、信号 $s \in S$ 和节点 $x \in h$，

$$q(s \mid a, x)\,\pi(a \mid h) = \mu((x, a, s) \mid x);$$

3. 对于每个信息集 $h \in H$、行动 $a \in A(h)$、信号 $s \in S$ 以及节点 x、$y \in h$，

$$\frac{\mu((x,a,s) \mid (x,a,s),(y,a,s))}{\mu((y,a,s) \mid (x,a,s),(y,a,s))} = \frac{q(s \mid a,x)}{q(s \mid a,y)} \frac{\mu(x \mid x,y)}{\mu(y \mid x,y)}.$$

在正文所考虑的博弈中，当发送者选择一项实验（初始实验或额外实验）时，行动是一项实验 $\sigma \in \Sigma$，相应的信号分布为 $\sigma(\cdot \mid x(\theta))$，其中 $x(\theta)$ 是节点 x 处的实际状态。当发送者选择要揭示的额外实验，或接收者选择行动时，相应的信号是确定性的，并且只导向一个后继节点。

引理 2. 设 (ν, π) 是一个完美贝叶斯均衡，且 μ 是相应的条件信念。对于任意接收者信息集 h，

1. 如果 $h \in \mathcal{H}_k$，那么 $\mu(\{t \geq k-1\} \mid h) = 1$。

2. $\mu(\{\theta \in \text{支撑集}(\widehat{\pi}_h)\} \mid h) = 1$。

3. 如果 $n = \text{最大值支撑集}(p)$ 且 $h \in \mathcal{H}_{n+1}$，那么 $\mu(\theta \mid h) = \widehat{\pi}_h(\theta)$ 对每个 $\theta \in \Theta$ 都成立。

证明。假设 $h \in \mathcal{H}_k$，那么对于每个 $x \in h$，相应的发送者类型 $t(x) \geq k-1$，因此陈述 1 成立。

假设 $\theta \notin \text{supp}(\widehat{\pi}_h)$。对于每个作为信息集 h前序的信息集 h'，相应的发送者后验信念也会将 θ的概率赋为 0。因此，根据完美贝叶斯均衡定义中的条件 3，接收者在 h 处的后验信念也必须将 θ 的概率赋为 0。

如果 $n = \max\{\text{supp}(p)\}$ 且 $h \in \mathcal{H}_{n+1}$，那么对于每个 $x \in h$，相应的发送者类型$t(x) = n$，因此接收者的后验信念为 $\mu(t = n \mid h) = 1$。这也意味着发送者必须已披露其进行的所有额外实验。根据完美贝叶斯均衡定义中的条件 3，接收者的后验信念就是接收者的朴素信念 $\widehat{\pi}_h$。

□

B 假设 1

令 $\Pi^* \subseteq \Delta(\Theta)$ 表示由 σ^* 诱导的后验信念集合。对于任意 $\pi \in \Delta(\Theta)$ 和$\varepsilon > 0$，令 $a^*(\pi)$ 表示信念为 π 时发送者偏好的接收者最佳回应，并令

$B(\pi,\ \varepsilon)$ 表示与 π 的距离小于 ε 的信念集合。以下结果展示了假设 1 的主要含义：存在一个接近 σ^* 的实验序列，其中每个实验几乎必然都会产生严格的接收者激励。[13] Lipnowski、Ravid 和 Shishkin (2025) 的引理 3 给出了一个相关结果。

引理 3。 对于任意 $\varepsilon > 0$，存在一个实验 σ_ε^*，使得对于每个 $\pi \in \Pi^*$，存在 $\pi_\varepsilon(\pi) \in B(\pi, \varepsilon)$，该分布由 σ_ε^* 诱导，且当接收者关于状态的信念为 $\pi_\varepsilon(\pi)$时，$a^*(\pi)$ 对接收者而言是严格最优的，并且 σ_ε^* 为 $\Pi_\varepsilon^* \equiv \{\pi_\varepsilon(\pi)\}_{\pi\in\Pi^*}$中的信念赋予的概率大于 $1-\varepsilon$。

证明。 由于 Θ 是有限的，且 $\Delta(\Theta)$ 是 $\mathbb{R}^{|\Theta|}$中的一个单纯形，因此根据 Carath´eodory 定理，任何均值为 π_0 的后验信念可行分布，都可以在不改变均值或发送者的期望收益的情况下，被替换为另一个支撑大小至多为 $|\Theta|$ 的分布。因此，不失一般性地，我们可以假设最优实验 σ^* 具有有限支撑 $\Pi^* = \{\pi^1, \ldots, \pi^m\}$，其中 $m \le |\Theta|$、$\sigma^*[\pi^j] > 0$和 $\sum_{j=1}^m \sigma^*[\pi^j]\pi^j = \pi_0$。

对每个 $j \in \{1, \ldots, m\}$，令 $a_j \equiv a^*(\pi^j)$。根据假设1，对于每个 j，存在$\tilde{\pi}^j \in \Delta(\Theta)$，使得 $\text{supp}(\tilde{\pi}^j) \subseteq \text{supp}(\pi^j)$，且 a_j 是 $\tilde{\pi}^j$上的唯一最佳响应。因此，根据期望的线性性，对于每个 $\delta > 0$ 和每个 j，存在 $\pi_\varepsilon(\pi^j) \in B(\pi^j, \delta)$，使得 a_j 是 $\pi_\varepsilon(\pi^j)$处的唯一最佳响应，且 $\text{supp}(\pi_\varepsilon(\pi^j)) \subseteq \text{supp}(\pi^j)$。

接下来，我们针对每个 j 构造一个定义在 $\Delta(\Theta)$ 上的概率测度 τ_j，使其 (i) 均值为 π^j，且 (ii) 几乎全部概率质量都集中在 $\pi_\varepsilon(\pi^j)$ 上。令 $\ell \equiv \min\left\{\pi^j(\theta) : j = 1, \ldots\ldots, m, \pi^j(\theta) > 0\right\}$，由于 Π^* 是有限的，且每个 π^j 都是概率向量，因此该值严格为正。固定 $\delta \le \frac{\varepsilon\ell}{2} < \varepsilon$ 和 $\eta \equiv \varepsilon \wedge \frac{1}{2}$。

对每个 j，定义向量 $\nu^j \in \mathbb{R}^{|\Theta|}$ 如下：$\nu^j(\theta) \equiv \frac{\pi^j(\theta) - (1-\eta)\pi_\varepsilon(\pi^j)(\theta)}{\eta}$ 对每个 $\theta \in \Theta$均成立。注意，

$$\sum_\theta \nu^j(\theta) = \frac{\sum_\theta \pi^j(\theta) - (1-\eta)\sum_\theta \pi_\varepsilon(\pi^j)(\theta)}{\eta} = \frac{1-(1-\eta)}{\eta} = 1$$

要证明 ν^j 是 非负，注意，如果 $\pi^j(\theta) = 0$，那么 $\pi_\varepsilon(\pi^j)(\theta) = 0$，因为 $\pi_\varepsilon(\pi^j) \in F_j$ ，

[13]为简便起见，本节仍将 σ^* 视为唯一最优实验（假设 2）。引理 3 的证明主要思路并不依赖这一额外假设。

因此 $\nu^j(\theta) = 0$。如果 $\pi^j(\theta) > 0$，那么 $\pi_\varepsilon(\pi^j)(\theta) \le \pi^j(\theta) + \delta$，因为 $\|\pi_\varepsilon(\pi^j) - \pi^j\| < \delta$。因此

$$\begin{aligned}\pi^j(\theta) - (1-\eta)\,\pi_\varepsilon(\pi^j)(\theta) &\ge \pi^j(\theta) - (1-\eta)\,(\pi^j(\theta) + \delta)\\ &= \eta\,\pi^j(\theta) - (1-\eta)\,\delta.\end{aligned}$$

利用 $\pi^j(\theta) \ge \ell$ 和 $\delta \le \varepsilon\ell/2$，再结合 $\eta \ge \varepsilon/2$（因为 $\eta = \varepsilon \wedge \frac{1}{2}$），可得 $\eta\,\pi^j(\theta) \ge \eta\,\ell \ge \frac{\varepsilon\ell}{2} \ge \delta$，因此 $\eta\,\pi^j(\theta) - (1-\eta)\,\delta \ge \delta - (1-\eta)\,\delta \ge 0$. 所以，对于所有 θ，都有 $\nu^j(\theta) \ge 0$，因此 $\nu^j \in \Delta(\Theta)$。

现在定义 τ_j 为两点分布 $\tau_j \equiv (1-\eta)\,\delta_{\pi_\varepsilon(\pi^j)} + \eta\,\delta_{\nu^j}$. 根据构造，$\mathbb{E}_{\pi\sim\tau_j}[\pi] = (1-\eta)\,\pi_\varepsilon(\pi^j) + \eta\,\nu^j = \pi^j$，并且 $\tau_j[\pi_\varepsilon(\pi^j)] = 1 - \eta \ge 1 - \varepsilon$.

定义一个新的后验信念分布 $\sigma^*_\varepsilon[\cdot] \equiv \sum_{j=1}^m \sigma^*[\pi^j]\,\tau_j[\cdot]$：首先根据原分布 σ^* 抽取 π^j，然后在给定 π^j 的条件下，从 τ_j 抽取实际实现的后验信念。由于 $\mathbb{E}_{\pi\sim\tau_j}[\pi] = \pi^j$，$\mathbb{E}_{\pi\sim\sigma^*_\varepsilon}[\pi] = \sum_{j=1}^m \sigma^*[\pi^j]\,\mathbb{E}_{\pi\sim\tau_j}[\pi] = \sum_{j=1}^m \sigma^*[\pi^j]\pi^j = \mathbb{E}_{\pi\sim\sigma^*}[\pi] = \pi_0$. 因此，$\sigma^*_\varepsilon \in \Sigma[\pi_0]$ 是贝叶斯合理的。令 $\Pi^*_\varepsilon \equiv \{\pi_\varepsilon(\pi^j)\colon\ j = 1, \dots, m\}$。对于每个 j，$\sigma^*_\varepsilon[\pi_\varepsilon(\pi^j)] = \sigma^*[\pi^j] \cdot \tau_j[\pi_\varepsilon(\pi^j)] = \sigma^*[\pi^j](1-\eta)$。因此，$\sigma^*_\varepsilon[\Pi^*_\varepsilon] = \sum_{j=1}^m \sigma^*[\pi^j](1-\eta) = 1 - \eta \ge 1 - \varepsilon$. 根据构造，接收者在 $\pi_\varepsilon(\pi^j)$ 处的唯一最佳响应是 $a^*(\pi^j)$，且根据构造，$\pi_\varepsilon(\pi^j) \in B(\pi^j, \delta) \subseteq B(\pi^j, \varepsilon)$，因为 $\delta \le \frac{\varepsilon\ell}{2} < \varepsilon$。因此，我们构造出了一个满足所需性质的、关于后验信念的贝叶斯合理有限支撑分布 σ^*_ε。

□

C 命题的证明1

在任何均衡中，发送者的策略都会诱导出接收者关于 θ 的后验信念分布，其期望值为 π_0。因此，发送者的均衡收益不高于其在辅助博弈中的收益；在该博弈中，发送者可以选择接收者关于 θ 的任意后验信念分布，但须满足期望值等于π_0这一约束。根据定义，这个上界为 $V(\pi_0)$。发送者还可以选择偏离：通过选择一个能完全揭示 θ 的初始实验，并且无论实现状态 θ如何，都不进行额外实验。在这种偏离下，对每个 $\theta \in \Theta$，当实现状态为 θ时，发送者获得收益 $u(\theta)$。因此，发送者通过这种偏离获得的期望收益为 $V(\pi_0)$，这意味着其均衡收益至少为 $V(\pi_0)$。

为证明当 p 是退化的时，发送者在所有均衡中的收益均为 $V(\pi_0)$，令 $u(\pi) \equiv \min_{a \in A^*(\pi)} u^s(a)$ 对每个 $\pi \in \Delta(\Theta)$ 均成立，并令

$$\overline{V}^{\sigma} \equiv \sum_{\pi \in \Delta(\Theta)} \sigma[\pi]\overline{u}(\pi) \quad \text{and} \quad \underline{V}^{\sigma} \equiv \sum_{\pi \in \Delta(\Theta)} \sigma[\pi]\underline{u}(\pi),$$

当接收者的信息结构为 σ 时，在分别倾向于有利于和不利于发送者的平局裁决规则下，发送者的期望收益为 Θ 和 A。由于 $_{\sigma \in \Sigma} V^{\sigma}$ 均为有限集，最大值点集 σ^* 非空。回顾可知，$V(\pi_0) = V^{\sigma^*}$ 是最优实验。根据定义，$V(\pi_0) = V^{\sigma^*}$。根据引理3，对于每个 $\varepsilon > 0$，存在一个实验 $\underline{\sigma}^{\varepsilon}$，使得 $V^{\sigma^{\varepsilon}} > V^{\sigma^*} - \varepsilon$。

假设 p 是退化的，并以概率 1 赋予类型 t^*。固定任意均衡和 $\varepsilon > 0$。考虑发送者在以下偏离下的收益：进行一个满足 $V^{\sigma^{\varepsilon}} > V^{\sigma^*} - \varepsilon = V(\pi_0) - \varepsilon$ 的初始实验 σ^{ε}。随后，无论初始实验的实现结果如何，发送者都进行 t^* 个额外实验，且这些实验均为无信息的。然后，他们向接收者披露所有这些额外实验的结果。在观察到 t^* 个额外结果后，接收者的后验信念以概率 1 认为发送者会披露所有结果，这意味着其关于 θ 的后验信念必须与其朴素信念一致。这意味着，在这种偏离下，发送者的收益至少为 $V(\pi_0) - \varepsilon$。由于 $\varepsilon > 0$ 可以是任意的，因此每个均衡中的发送者收益至少为 $V(\pi_0)$。

D 引理证明 1

由于每个退化信念都是可信的，且假设 σ^* 是唯一最优实验，因此假设 σ^* 诱导了不可信信念，意味着 $V(\pi_0) > V(\pi_0)$。令 $\pi \in \Delta(\Theta)$ 表示由 σ^* 诱导的不可信信念。根据定义，存在 $\theta \in \text{supp}(\pi)$，使得 $u(\theta) > u(\pi)$。回忆一下，$B(\pi, d)$ 表示以 π 为中心、半径为 $d > 0$ 的开球。对于每个 $\sigma \in \Sigma[\pi_0]$，记给定状态 θ 时信念的条件分布为 σ^{θ}，其中 $\sigma^{\theta}[B] \equiv \int_B {}^{\pi}{}_{(\theta)}\, {}_{\pi_0(\theta)} d\sigma[\pi]$ 对每个博雷尔可测集 $B \subseteq \Delta(\Theta)$ 都成立。

引理 4。假设 r 满足：对于每个 $\pi \in B(\underline{\pi}, r)$，$\pi(\theta) > \pi(\theta)/2$。则存在 η、$\varepsilon > 0$，使得对于每个 $\sigma \in \Sigma[\pi_0]$，若 $\mathbb{E}_{\pi\sim\sigma} u(\pi) > V(\pi_0) - \eta$，则 $\sigma^\theta[B(\pi, r)] \geq \varepsilon$。

引理 4 的证明。令 $m \equiv \sigma^*[B(\underline{\pi}, r)] > 0$，并令 $\mathcal{C} := \{\sigma \in \Sigma[\pi_0] : \sigma[B(\underline{\pi}, r)] \leq m/2\}$。由于 $B(\pi, r)$ 是一个开集，根据波特曼托定理，将 σ 映射到 $\sigma[B(\pi, r)]$ 的函数是下半连续的，因此下水平集 $\{\sigma \in \Delta(\Delta(\Theta)) : \sigma[B(\pi, r)] \leq m/2\}$ 是闭集。由于 Θ 是有限的，$\Sigma[\pi_0]$ 是紧的，因此 $\mathcal{C}$ 是紧集的闭子集，因而也是紧的。

根据定义，$\sigma^* \notin \mathcal{C}$。由于 σ^* 是唯一最优实验（假设 2），$\sup_{\sigma\in\mathcal{C}} \mathbb{E}_{\pi\sim\sigma}(u(\pi)) < V(\pi_0)$。令 $\eta \equiv \min\{V(\pi_0) - \sup_{\sigma\in\mathcal{C}} \mathbb{E}_\sigma(u), \overline{\pi}/2\} > 0$，并令 $\varepsilon \equiv \pi(\theta) 2\pi_0(\theta)\eta > 0$。由此可知，对于任意 $\sigma \in \Sigma[\pi_0]$，只要 $\mathbb{E}_{\pi\sim\sigma}(u(\pi)) > V(\pi_0) - \eta$，$\sigma \notin \mathcal{C}$，因此 $\sigma[B(\pi, r)] > m/2 \geq \eta$。根据定义，

$$\sigma^\theta[B(\underline{\pi}, r)] = \int_{B(\underline{\pi},r)} \frac{\pi(\theta)}{\pi_0(\theta)} d\sigma[\pi] \geq \frac{\underline{\pi}(\theta)}{2\pi_0(\theta)} \sigma[B(\underline{\pi}, r)] \geq \frac{\underline{\pi}(\theta)}{2\pi_0(\theta)} \eta = \varepsilon,$$

其中，第一个不等式由以下假设得出：对于每个 $\pi \in B(\pi, r)$，$\pi(\theta) > \pi(\theta)/2$。

□

令 $\overline{v}$ 和 $\underline{v}$ 分别表示发送者可行收益的最高值和最低值。由于 $u(\pi)$ 关于 π 是上半连续函数，因此存在 $r > 0$，使得对于每个 $\pi \in B(\pi, r)$，都有 $u(\theta) > u(\pi)$ 且 $\pi(\theta) > \pi(\theta)/2 \geq 0$。根据引理 4，存在 $\eta, \varepsilon > 0$，使得对于每个 $\sigma \in \Sigma[\pi_0]$，若 $\mathbb{E}_{\pi\sim\sigma} u(\pi) > V(\pi_0) - \eta$，则 $\sigma^\theta[B(\pi, r)] \geq \varepsilon$。固定这一组 r, η, ε，并选择满足以下条件的 $N \in \mathbb{N}$

$$N > \frac{3(\overline{v} - \underline{v})}{2\varepsilon\pi_0(\theta)(u(\theta) - \overline{u}(\underline{\pi}))} + \frac{2}{\varepsilon}.$$

在类型分布集合 $\Delta(\mathbb{N})$ 上赋予积拓扑。对于任意属于

$$U \equiv \left\{ p \in \Delta(\mathbb{N}) : |p(t) - \tfrac{1}{N}| < \tfrac{1}{2N} \, \forall t = 0, \ldots, N-1, \text{ and } p(\{t \geq N\}) < \tfrac{1}{2N} \right\},$$

我们证明，任何均衡中的发送者期望收益都不超过 $V(\pi_0) - \eta$。

反设存在一个均衡，使得发送者收益严格大于 $V(\pi_0) - \eta$。令 $\widetilde{\sigma}$ 表示接收者后验信念的均衡分布。

对于任意 $t \in \{0, 1, ..., N\}$，选取任意一种在均衡中以正概率被类型 t 采用的纯策略，并令 $\mathcal{H}_t^*$ 表示如下信息集的集合：对于每个$h \in \mathcal{H}_t^*$，在类型 t采用的该纯策略下，(i) h 以正概率到达，且该类型 t 此后不再进行额外实验；(ii) 在 h 披露额外实验结果的均衡概率会诱导出一个属于 $B(\pi, r)$ 的接收者后验信念。

对于每个 $t \geq 0$，考虑类型 $t+1$偏离到以下策略时的收益：采用与类型 t相同的策略，但对属于 $\mathcal{H}_t^*$的历史除外。在每个 $h \in \mathcal{H}_t^*$，进行一次能够完全揭示状态的额外实验，并且当且仅当状态为 θ时披露该实验的结果。令 V_{t+1} 表示类型$t+1$ 的均衡收益，并令 $\widetilde{\sigma}_{t+1}^{\theta}[\cdot]$ 表示在状态为 θ时，类型 $t+1$采用其均衡策略所诱导的接收者后验信念分布。类型 $t+1$通过上述偏离获得的收益至少为

$$V_t + \widetilde{\sigma}_t^{\theta}[B(\underline{\pi}, r)]\pi_0(\theta)(u(\theta) - \overline{u}(\underline{\pi})),$$

因为它们会诱使接收者采取与类型 t 在均衡中以正概率使用的某一纯策略相同的行动；但当状态为 θ 且类型 t 到达属于 $\mathcal{H}_t^*$ 的信息集时，类型 $t+1$ 的偏离会诱使接收者采取针对状态 θ 的最优行动。因此，对于每个 $t \geq 0$，我们有

$$V_{t+1} - V_t \geq \widetilde{\sigma}_t^{\theta}[B(\underline{\pi}, r)]\pi_0(\theta)\Big(u(\theta) - \overline{u}(\underline{\pi})\Big).$$

由此可得

$$V_{N-1} - V_0 \geq \left(\sum_{t=0}^{N-2} \widetilde{\sigma}_t^{\theta}[B(\underline{\pi}, r)]\right) \pi_0(\theta)\Big(u(\theta) - \overline{u}(\underline{\pi})\Big).$$

由于对每个 $t = 0, ..., N-1$，$|p(t) - \frac{1}{N}| \leq \frac{1}{2N}$，我们有

$$(\tfrac{1}{N} + \tfrac{1}{2N}) \sum_{t=0}^{N-2} \tilde{\sigma}_t^{\theta}[B(\underline{\pi}, r)] \geq \sum_{t=0}^{N-2} p(t)\widetilde{\sigma}_t^{\theta}[B(\underline{\pi}, r)] = \widetilde{\sigma}^{\theta}[B(\underline{\pi}, r)] - \sum_{t=N-1}^{\infty} p(t)\widetilde{\sigma}_t^{\theta}[B(\underline{\pi}, r)]$$

根据引理 4，我们有

$$\widetilde{\sigma}^{\theta}[B(\underline{\pi},r)]-\sum_{t=N-1}^{\infty}p(t)\widetilde{\sigma}_{t}^{\theta}[B(\underline{\pi},r)]\geq\varepsilon-p(N-1)-p(\{t\geq N\})\geq\varepsilon-\tfrac{2}{N},$$

其中，最后一个不等式由 $|p(N-1)-\frac{1}{N}|\leq\frac{1}{2N}$ 和 $p(\{t\geq N\})\leq\frac{1}{2N}$ 得出。因此，

$$V_{N-1}-V_{0}\geq(\tfrac{2}{3}N\varepsilon-\tfrac{4}{3})\pi_{0}(\theta)(u(\theta)-\overline{u}(\underline{\pi})).$$

这与 N 的构造一起导致矛盾。

E 定理2的证明

由于 $u(\theta)>\max^{\pi\in\Pi_{\theta}^{*}}u(\pi)$，且最佳回应对应是上半连续的，因此存在 $\varepsilon>0$，使得对于任意满足 π' 与 Π_{θ}^{*} 之间的豪斯多夫距离小于 ε 的 π'，都有 $u(\pi')\leq\max^{\pi\in\Pi_{\theta}^{*}}u(\pi)$。由于最优实验是唯一的，且发送者收益有界，我们知道存在 $\lambda_{0}>0$，使得对于任意 $\eta>0$，以及任意发送者收益大于 $V(\pi_{0})-\eta$ 的均衡，条件于每个 $\theta\in\Theta$，接收者以至少 $1-\lambda_{0}\eta$ 的概率到达信息集，而其后验信念与 Π_{θ}^{*} 之间的豪斯多夫距离小于 η。令 v 和 v 分别表示发送者最高和最低的可行收益。

令 $\lambda\equiv2\lambda_{0}$ 且

$$n\equiv\left\lceil\frac{2(\overline{v}-\underline{v})}{\pi_{0}(\theta)\left\{u(\theta)-\max\limits_{\pi\in\Pi_{\theta}^{*}}\overline{u}(\pi)\right\}}\right\rceil+1,$$

并任取满足以下条件的 η

$$\eta<\min\left\{\varepsilon,\frac{1}{\lambda n}\right\}.$$

反设 p 赋予至少 n种类型的概率大于 $\lambda\cdot\eta$，但存在一个均衡，使发送者收益大于 $V(\pi_{0})-\eta$。根据构造，以 θ 为条件时，接收者的后验信念与集合 Π_{θ}^{*} 的豪斯多夫距离小于 η 的概率大于 $1-\lambda\eta/2$。令 $t_{1}<\ldots<t_{n}$ 表示出现概率至少为 $\lambda\eta$的 n种类型。令 $\Pr(\cdot\mid\theta,t)$ 表示

在状态为 θ时，类型为 t 的均衡策略所对应的概率测度。我们用 D 表示诱导的接收者后验信念与 Π^*_θ之间的豪斯多夫距离，它是一个随机变量。对各类型取期望，可得 $\mathbb{E}_{t\sim p}[\Pr(D \geq \eta \mid \theta, t)] < \lambda\eta/2$。由马尔可夫不等式可知，$p(\{t \in \Delta(\mathbb{N}) : \Pr(D \geq \eta \mid \theta, t) \geq 1/2\}) < \lambda\eta$。由于每种选定类型$t_i$ 的概率至少为 $\lambda\eta$，它们都不可能属于该集合，因此对每个 t_i ，都有 $\Pr(D < \eta \mid \theta, t_i) > 1/2$。对每个 $i \in \{1, 2, ..., n\}$，任选一种在均衡中由类型 t_i 以正概率采用的纯策略，并令 $\mathcal{H}^*_i$ 表示满足下述条件的发送者历史集合：对每个 $h \in \mathcal{H}^*_i$，在类型为 t_i时采用该纯策略，将以正概率到达 h，此后不再进行任何额外实验；并且在 h 披露额外实验结果的均衡概率，将使接收者的后验信念与 Π^*_θ之间的豪斯多夫距离小于 η。

对于每个 $i \geq 2$，考虑类型 t_i偏离到以下策略时的收益：除了属于 $\mathcal{H}^*_{i-1}$ 的历史之外，采用与类型 t_{i-1} 相同的策略。在每个 $h \in \mathcal{H}^*_{i-1}$，进行一次能够完全揭示状态的额外实验，并且当且仅当状态为 θ 时披露该实验的结果。令 V_i 表示类型 t_i的均衡收益。类型 t_i通过上述偏离获得的收益至少为

$$V_{i-1} + \frac{1}{2}\pi_0(\theta)\Big\{u(\theta) - \max_{\pi\in\Pi^*_\theta} \overline{u}(\pi)\Big\},$$

因为他们会诱导出与类型 t_{i-1} 在均衡中以正概率采用的某个纯策略相同的接收者行动，但当状态为 θ 且类型 t_{i-1} 到达历史 $\mathcal{H}^*_{i-1}$时，他们会诱导出接收者行动 $a^*(\theta)$。因此，对于每个 $i \geq 2$，我们有

$$V_i - V_{i-1} \geq \frac{1}{2}\pi_0(\theta)\Big\{u(\theta) - \max_{\pi\in\Pi^*_\theta} \overline{u}(\pi)\Big\}.$$

这与 n 的构造一起导出了矛盾。

F 定理 2 中的要求2 并非冗余

我们用一个示例说明，定理 2 中的要求无法弱化为引理 1中的要求。设 $A \equiv \{0, 2, 3\}$，$\Theta \equiv \{\theta_0, \theta_1\}$，以及 $t \in \{0, 1\}$。接收者的先验信念

赋予状态 θ_1 的概率为 π_0。发送者的收益等于 a。若 $\pi \in [0, 1/3]$，接收者的最优行动为 0；若 $\pi \in [1/3, 2/3]$，则为 2；若 $\pi \in [2/3, 1]$，则为 3。先验状态分布为 $\pi_0 = 1/2$，因此最优实验诱导出信念 $1/3$ 和 $2/3$。信念 $1/3$ 不可信，但发送者在状态 θ_1 下的收益与其诱导出信念 $2/3$ 时的收益相同。设先验类型分布 $p \in \Delta(\mathbb{N})$ 由下式给出：对所有 $t \in \mathbb{N} \setminus \{0\}$，$p(0) = 1/3$，且 $p(t) = \frac{1}{3}\frac{1}{2^{t-1}}$。

我们将证明，这个示例存在一个均衡，其中 (i) 在未披露额外结果的每个信息集上，接收者的后验信念为 $1/3$；在披露一个额外结果的每个均衡路径信息集上，接收者的后验信念为 $2/3$；(ii) 发送者进行无信息初始实验；(iii) 如果发送者实现的类型 t 至少为 1，发送者将进行 t 个额外实验，每个实验都有两个信号实现 $\{s, s\}$，分别为 $\sigma(s|\theta_0)$ 和 $\sigma(s|\theta_1)$，且只披露第一个实现的、满足 s 的结果；(iv) 在每个具有非退化朴素信念的非均衡路径信息集上，接收者的后验信念为 0。

为证明这样的均衡存在，我们首先验证均衡路径上的接收者后验信念与贝叶斯法则一致。观察到一个额外结果 (σ, s) 后，接收者的后验信念为

$$\frac{\pi_0(1-p(0))}{\sum_{t=1}^{\infty} p(t)(1-(1-\pi_0)(\frac{2}{3})^t)} = \frac{\frac{1}{3}}{\frac{1}{3}\sum_{t=1}^{\infty}\frac{1}{2^{t-1}}(1-\frac{1}{2}(\frac{2}{3})^t)} = \frac{2}{3}.$$

在观察到没有额外结果后，接收者的后验信念为

$$\frac{\pi_0 p(0)}{p(0)+\sum_{t=1}^{\infty} p(t)(1-\pi_0)(\frac{2}{3})^t} = \frac{\frac{1}{6}}{\frac{1}{3}+\frac{1}{3}\sum_{t=1}^{\infty}\frac{1}{2^t}(\frac{2}{3})^t} = \frac{1}{3}.$$

接下来，我们检验不存在任何发送者类型有动机偏离。对于类型 t，均衡路径上的收益可以递归计算：首先进行实验 σ，如果 s 实现，则披露结果 (σ, s)；否则，继续进行实验，直到在剩余的 $t-1$ 次实验中实现信号 s。s 以概率 $2/3$ 实现，从而获得收益 $u(2/3)$；否则，发送者获得续续收益 $V_{t-1} = (1-(2/3)^{t-1})u(2/3) + (2/3)^{t-1}u(1/3)$，即在给定 $\theta = \theta_0$ 的情况下重复 σ $t-1$ 次所得到的期望收益。因此，类型 t 的均衡路径上的收益为 $\frac{2}{3}u(2/3) + \frac{1}{3}V_{t-1}$。对于类型 t 而言，最有利的偏离是首先进行一个完全信息附加实验，披露

如果 $\theta = \theta_1$实现，则披露结果；否则重复实验 σ $t-1$ 次，并且只披露第一个实现且为 s 的结果。我们同样可以递归计算收益，即 $\frac{1}{2}u(1) + \frac{1}{2}V_{t-1}$。均衡路径上的收益高于最佳偏离，因为 $V_{t-1} < 3 = u(2/3)$。

G 定理3的证明

当最优实验 σ^*（根据假设 2，它是唯一的）完全揭示 θ 时，我们有$V(\pi_0) = V(\pi_0)$；此时，命题 1 将意味着存在一个均衡，其收益为 $V(\pi_0)$。接下来，我们聚焦于 σ^* 未能完全揭示 θ 的情形。根据引理 3，对于每个 $\delta > 0$，都存在 $\varepsilon > 0$ 和一个实验 σ_ε^*，使得对于每个$\pi \in \Pi^*$，都存在由 σ_ε^* 导致的 $\pi_\varepsilon(\pi) \in B(\pi, \varepsilon)$，并且当接收者关于状态的信念为 $\pi_\varepsilon(\pi)$ 时， $a^*(\pi)$ 对接收者而言是严格最优的，同时 σ_ε^* 将超过 $1-\varepsilon$ 的概率赋予 $\Pi_\varepsilon^* = \{\pi_\varepsilon(\pi)\}_{\pi\in\Pi^*}$ 中的信念。因此，发送者承诺采用任何此类 σ_ε^* 所获得的期望收益至少为 $V(\pi_0) - \delta$。

固定一个类型分布 $p \in \Delta(\mathbb{N})$，它以概率 $p(n) > 1-\varepsilon$ 赋予某个 $n \in \mathbb{N}$。我们证明存在一个完美贝叶斯均衡，使发送者收益高于 $V(\pi_0) - \delta$。如果 supp(p) 有上界，令 $N \equiv \max \text{supp}(p)$ 表示可能的最高发送者类型。我们着重讨论 $n < N$ 或 supp(p) 无上界的情形（不存在这样的 N），因为如果 $n = N$，运用与命题 1 类似的论证，类型为 n 的发送者可以确保获得接近其承诺收益的收益。由此可知，发送者的事前收益也接近承诺收益，因为 $p(n) \approx 1$。

如果 $n < N$ 或 supp(p) 无上界，我们按如下方式构造一个均衡：

- 发送者进行一次初始实验 σ_ε^*。接收者观察到该初始实验的结果后所持有的朴素信念称为其中期信念。

- 在 Π_ε^* 中的任何可信中期信念之后，除最高类型 N （如果存在）外，发送者的所有类型在均衡中都不进行额外实验。

- 在 Π_ε^* 中出现任何不可信中期信念后，类型 1 至 $n-1$ 会进行完全信息附加实验，并披露部分状态，同时隐瞒其他状态。

类型 n 会进行 n 无信息实验并披露所有结果。根据类型 0 与类型 1 至 $n-1$ 的相对概率，他们也可能模仿较低类型；我们在小节 G.1 中分别讨论两种情况。严格大于n 但不包括 N 的类型会进行 n 无信息实验，披露所有结果，然后再进行一次完全信息实验，并隐瞒不利的状态报告。

- 不在 Π_ε^* 中的中期信念在 σ_ε^* 下发生的概率小于 ε ，因此对发送者的事前期望收益影响可以忽略不计。

- 接收者在均衡路径上的后验信念与贝叶斯法则一致。在任意非均衡路径信息集 $h \notin \mathcal{H}_{N+1}$处，如果发送者揭示任何具有信息性并诱导出非退化的朴素信念 π的额外信号，那么接收者的后验信念会以概率 1 赋予某个满足下式的 $\theta \in$supp(π)

$$u(\theta) \leq u(\theta') \text{ for every } \theta' \in \text{supp}(\pi).$$

在该均衡中，在任意中期信念 $\pi \in \Pi_\varepsilon^*$之后，类型 n在均衡路径上的收益接近于 $u(\pi)$，因此发送者的事前期望收益接近于 $V(\pi_0)$，当 $p(n) \approx 1$时亦然。鉴于接收者在非均衡路径上的后验信念，不失一般性地，我们可以将注意力集中于发送者不偏离到其他初始实验的子博弈。在下面一段中，我们分析在 Π_ε^*中的可信中期信念之后的子博弈。在小节G.1中，我们考察在 Π_ε^* 中的不可信中期信念之后的子博弈，并描述策略组合的细节。

对于任意可信中期信念 $\pi \in \Pi_\varepsilon^*$，除最高类型 N （如果存在）外，发送者的所有类型在均衡中都不进行额外实验。鉴于接收者的均衡路径外信念，以及以下事实，这对发送者的所有类型而言都是最优的：(i) π 是可信的；(ii) 类型 N 的概率可以忽略不计，因此接收者在观察到没有进行额外实验时的后验信念接近 π ，且同样可信。

G.1 不可信中期信念

考虑中期信念 $\pi \in \Pi_\varepsilon^*$ 不可信之后的子博弈。令 supp(π) $\equiv\{\theta_1, ..., \theta_k\}$, $\pi_j \equiv \pi(\theta_j)$，且对 $j \in \{1, \ldots, k\}$ 有 $u_j \equiv u(\theta_j)$。不失一般性，我们

假设 $u_1 \geq u_2 \geq ... \geq u_k$。令 $u^* \equiv u^s(a^*(\pi))$。由于 $\pi \in \Pi^*_\varepsilon$ 不可信，且完全披露在 π 处并非最优，我们知道 $u_1 > u^* > u_k$。

令 u^*_i 由下列方程定义：

$$u^* = \sum_{j=1}^{i} \pi_j u_j + \Big(1 - \sum_{j=1}^{i} \pi_j\Big) u^*_i. \tag{2}$$

直观地说，u^*_i 是发送者在不披露时的收益，因此，只有当状态属于 u^* 时，他才会在获得 $\{\theta_1, ..., \theta_i\}$与披露状态之间无差异。根据定义，$u^*_k = +\infty$。发送者揭示索引低于 i 的状态，并隐瞒索引高于 i 的状态，这一激励要求 $u_i \geq u^*_i \geq u_{i+1}$。引理5 表明，这一激励约束等价于 u^*_i 在 i处达到局部最小值。

引理5。 $u_i \geq u^*_i \geq u_{i+1}$ 当且仅当$\min\{u^*_{i+1}, u^*_{i-1}\} \geq u^*_i$。

证明。将方程（2）应用于 i 和 $i+1$，可得：

$$\sum_{j=1}^{i} \pi_j u_j + \Big(1 - \sum_{j=1}^{i} \pi_j\Big) u^*_i = \sum_{j=1}^{i+1} \pi_j u_j + \Big(1 - \sum_{j=1}^{i+1} \pi_j\Big) u^*_{i+1}. \tag{3}$$

方程（3）等价于

$$\Big(1 - \sum_{j=1}^{i+1} \pi_j + \pi_{i+1}\Big) u^*_i = \pi_{i+1} u_{i+1} + \Big(1 - \sum_{j=1}^{i+1} \pi_j\Big) u^*_{i+1},$$

这可以改写为

$$\Big(1 - \sum_{j=1}^{i+1} \pi_j\Big)(u^*_i - u^*_{i+1}) = \pi_{i+1}(u_{i+1} - u^*_i). \tag{4}$$

这意味着 $u^*_{i+1} \geq u^*_i$当且仅当 $u^*_i \geq u_{i+1}$。代入 $\sum_{j=1}^{i+1} \pi_j = \sum_{j=1}^{i} \pi_j + \pi_{i+1}$并重新整理各项，我们可由（3）得到：

$$\Big(1 - \sum_{j=1}^{i} \pi_j\Big) u^*_i = \pi_{i+1} u_{i+1} + \Big(1 - \sum_{j=1}^{i} \pi_j - \pi_{i+1}\Big) u^*_{i+1},$$

由此可得

$$\Big(1-\sum_{j=1}^{i}\pi_j\Big)(u_i^*-u_{i+1}^*)=\pi_{i+1}(u_{i+1}-u_{i+1}^*). \tag{5}$$

这意味着 $u_{i+1}\geq u_{i+1}^*$ 当且仅当 $u_i^*\geq u_{i+1}^*$。类似的论证表明$u_{i-1}^*\geq u_i^* \iff u_i\geq u_i^*$.

□

引理 6. 存在 $i\in\{1,2,...,k\}$，使得 $u_i\geq u_i^*\geq u_{i+1}$。

证明。考虑两种情况。首先，如果 $u_1^*\geq u_2^*$，那么根据 $u_k^*=+\infty$ 可知存在 i，使得 $\min\{u_{i+1}^*,u_{i-1}^*\}\geq u_i^*$。引理 5 意味着，对于这样的 $u_i\geq u_i^*\geq u_{i+1}$ i。其次，如果 $u_2^*>u_1^*$，那么不等式（4）意味着 $u_1^*\geq u_2$。为了证明 $i=1$ 满足我们的要求，我们只需证明 $u_1\geq u_1^*$。反设$u_1<u_1^*$，那么鉴于 $u_1\pi_1+(1-\pi_1)\,u_1^*=u^*$，可得 $u_1<u^*<u_1^*$。这与 π 是不可信的这一推论相矛盾，即 $u_1>u^*$。

□

定义策略 s^i 为发送者纯策略：在中期信念π之后的子博弈中，发送者仅进行一次完全信息附加实验，并且当且仅当 $\theta\in\{\theta_1,...,\theta_i\}$时披露由此产生的结果。回想一下，类型 n 发生的概率接近 1。定义策略 s^{nd} 为一种发送者纯策略，在该策略下，发送者进行 n次无信息实验并披露所有结果。根据定义，策略 s^i 仅对至少为 1 的类型可行，而策略 s^{nd} 仅对至少为 n的类型可行。令$\mathbb{G}\subseteq\Delta(\Theta)\times\mathbb{R}$ 满足：当且仅当存在 $\alpha\in\Delta(A)$，使得 α 是后验信念 π 下接收者的最佳回应，且 $\mathbb{E}_{a\sim\alpha}[u^s(a)]=u$时，$(\pi,\ u)\in\mathbb{G}$。根据定义，$\mathbb{G}$ 是闭且连通的。

固定任意 i 满足 $u_i\geq u_i^*\geq u_{i+1}$的元素，该元素由引理6保证存在。令 π_*^i 表示接收者的后验信念，其向状态 θ_j 赋予概率 $\frac{\pi_j}{\sum_{\tau=i+1}^{k}\pi_\tau}$，对于 $j\in\{i+1,...,k\}$，并向其他状态赋予零概率。当 ε 足够小时，$\pi\in\Pi_\varepsilon^*$意味着 π 接近于发送者的最优实验 σ^*下的中期信念 π'，且 $a^*(\pi)=a^*(\pi')$。由于 σ^* 是最优实验，我们知道对于每一个

$a \in$ 最大值点集$_{a' \in A}\, u^r(\pi^i_*,\ a')$,

$$u^* > \sum_{j=1}^{i} \pi_j u_j + \Big(1 - \sum_{j=1}^{i} \pi_j\Big) u^s(a). \tag{6}$$

在中期信念 π下，令 π^i 表示在某一历史处接收者关于 θ 的后验信念。对于每个 $\theta \in \{\theta_1, ..., \theta_k\}$，该历史都由类型 0 引致；当且仅当 $\theta \in \{\theta_{i+1}, ..., \theta_k\}$时，该历史由类型 1 至 $n-1$引致；而类型 n 及以上则绝不会引致该历史。形式上，令 $p(t)$表示接收者的先验信念赋予类型 t的概率。于是，信念 π^i 定义如下:

$$\pi^i(\theta_j) \equiv \frac{p(0)\pi_j}{p(0) + \sum_{t=1}^{n-1} p(t) \cdot \sum_{s=i+1}^{k} \pi_s} \text{ for every } j \leq i$$

and

$$\pi^i(\theta_j) \equiv \frac{\sum_{t=0}^{n-1} p(t)\pi_j}{p(0) + \sum_{t=1}^{n-1} p(t) \cdot \sum_{s=i+1}^{k} \pi_s} \text{ for every } j \geq i+1.$$

根据接收者在 π^i处的最佳回应，我们考虑两种情形。

情形A： 假设存在 $a \in$ 最大值点集$_{a' \in A}\, u^r(\pi^i,\ a')$，使得

$$u^* \leq \sum_{j=1}^{i} \pi_j u_j + \Big(1 - \sum_{j=1}^{i} \pi_j\Big) u^s(a). \tag{7}$$

直观而言，这是类型 0 的概率相对于类型 1 至 $n-1$的概率较高的情形。(u_i^*2)中的定义意味着 $u^s(a) \geq u_i^*$。

在中期信念 π之后的均衡中，严格高于 n 的发送者类型（若最高类型 N 存在，则不包括该类型）进行 n 无信息实验，披露所有结果，然后再进行一次完全信息实验；当且仅当实现状态 $\theta = \theta_j$ 满足 $u_j > u^*$时，才披露该实验的结果。发送者类型 N 可以采用不同的策略，但该策略既不同于 s^i，也不同于 s^{nd}。发送者类型 1 至 $n-1$ 采用策略 s^i。

发送者类型 n 在策略 s^{nd} 和策略 s^i之间进行混合。如果类型 n 确定采用策略 s^{nd}，那么接收者在观察到没有额外实验时的后验信念为 π^i。如果类型 n 确定采用策略 s^i，由于这种情况发生的概率接近 1，那么接收者在观察到没有额外实验时的后验信念 $\widetilde{\pi}^i_*$ 接近 π^i_*，因此存在 $a \in$ 最大值点集$_{a' \in A}\, u^r(\widetilde{\pi}^i_*,\ a')$

满足（6）。由于图 $\mathbb{G}$ 是闭且连通的，类型 n 可以在策略s^i 和 s^{nd} 之间进行混合，使得在未披露额外结果时接收者的后验信念下，它们存在一个（可能是混合的）最佳回应 $\alpha \in \Delta（A）$，满足 $\mathbb{E}_{a\sim\alpha}[u^s（a）] = u_i^*$，从而使类型 n 在 s^i 和 s^{nd}之间无差异。

情形 B： 接下来，假设对于每个 $a \in$， $u^* > \sum_{j=1}^{i} \pi_j u_j + \left(1-\sum_{j=1}^{i} \pi_j\right) u^s(a)$， a'的最大值点集。直观地说，这种情形下类型 0 的概率相对于类型 1 到 $n-1$ 的概率较低。根据定义，$u^s(a) < u_i^*$。由于 i 的构造要求 $u_i^* \in [u_{i+1}, u_i]$，我们知道 $u^s(a) < u_i$。

在中期信念 π 之后的均衡路径上，严格大于 n 的发送者类型（如果最高类型 N 存在，则不包括该类型）将进行无信息实验，披露所有结果，然后再进行一次完全信息实验，并且仅当实现状态为满足 $u_j > u^*$ 的 $\theta = \theta_j$ 时，才披露该实验的结果。发送者类型 N 可以采用不同的策略，但该策略既不是 s^j，也不是 s^{nd}。发送者类型 1 到 $n-1$ 以某个 $j \geq i$ 的概率在策略 s^j 和策略 s^{j+1} 之间随机化。发送者类型 n 确定采用策略 s^{nd}。

我们证明，存在 $j \geq i$，以及类型 1 到 $n-1$ 在策略 s^j 与策略 s^{j+1} 之间的混合概率，使得在接收者观察到没有披露任何额外结果后形成的后验信念下，存在对接收者而言最优的 $\alpha \in \Delta(A)$，并且如果接收者在不披露时采取行动 α，那么以下两个激励约束均得到满足。首先，

- 当策略 s^j 和 s^{j+1} 都以正概率实施时，对于每个 $\tau \in \{1, 2, ..., k\}$，发送者从策略 s^j 和 s^{j+1} 获得的收益均不低于从策略 s^τ 获得的收益。

- 当只有策略 s^j 以正概率实施时，对于每个 $\tau \in \{1, 2, ..., k\}$，发送者从策略 s^j 获得的收益不低于从策略 s^τ 获得的收益。

这确保发送者类型 1 至 $n-1$ 没有动机偏离到其他策略。其次，发送者从策略 s^j 和 s^{j+1} 中获得的期望收益不超过 s^{nd}。这确保发送者类型 n 没有动机去模仿类型 1 至 $n-1$。

一个披露规则 是从 $\{\theta_1, ..., \theta_k\}$ 到各个状态被披露时所对应概率的映射。一个单调披露规则 是指：对于每个以正概率被披露的 θ_j，所有满足 $\tau < j$ 的状态 θ_τ 都以概率 1 被披露。根据定义，每个单调披露规则都可以由其披露概率唯一确定，记为 q。上文所描述的类型 1 至 $n-1$ 的策略都映射到单调披露规则。

我们证明，存在一种单调披露规则，使第一项激励约束得到满足。令 $V(q) \subseteq \mathbb{R}$ 表示：当接收者认为发送者类型 1 到 $n-1$ 采用披露概率为 q的单调披露规则时，发送者不披露额外信号所获得的期望收益集合。根据定义，$u^* \in V(1)$，因为当类型 1 到 $n-1$的披露概率为 1 时，只有类型 0 不披露额外信号。令 $\mathbb{V} \subseteq [0,1] \times \mathbb{R}$ 满足：$(q, v) \in \mathbb{V}$ 当且仅当 $v \in V(q)$。令 $\mathbb{U} \subseteq [0,1] \times \mathbb{R}$ 满足：(i) 对于每个可写为 $q = \sum_{j=1}^{\tau} \pi_j$（其中 τ为某个整数）的 q，有 $(q, v) \in \mathbb{U}$当且仅当 $v \in [u_{\tau+1}, u_\tau]$；(ii) 对于每个满足 $\sum_{j=1}^{\tau} \pi_j < q < \sum_{j=1}^{\tau+1} \pi_j$的 q，有 $(q, v) \in \mathbb{U}$ 当且仅当 $v = u_{\tau+1}$。$\mathbb{V}$ 和 $\mathbb{U}$ 均连续，而且，图像 $\mathbb{U}$ 关于 q 递减，且经过 $(1, u_k) \in \mathbb{U}$。本情形的假设要求，当 $q = q^* \equiv \sum_{j=1}^{i} \pi_j$时，$V(q^*)$中的每个收益都小于 u_i。由于 $u_k < u^*$，$\mathbb{V}$ 与 $\mathbb{U}$之间存在交点 $(\hat{q}, \hat{v})$，其中 $\hat{q} \geq q^*$。根据定义，任何披露概率为 q 的单调披露规则，以及导致不披露收益 v 的接收者行动，都将满足第一项激励约束。

在最后一步中，我们验证交点 $(\hat{q}, \hat{v})$ 也满足第二个激励约束。如果交点位于 $\hat{q} = q^*$，那么第二个激励约束显然得到满足，因为对于每个 $a \in \arg\max_{a' \in A} u^r(\pi^i, a')$，都有 $u_i^* \in [u_{i+1}, u_i]$ and $u^* > \sum_{j=1}^{i} \pi_j u_j + \left(1 - \sum_{j=1}^{i} \pi_j\right) u^s(a)$。下文中，我们聚焦于交点位于 $\hat{q} > q^*$ 的情形。对于每个 $\tau \in \{1, 2, ..., k\}$ and $\beta \in [0,1]$，令 $u^*(\tau, \beta)$ 定义为

$$u^*(\tau, \beta) = \frac{u^* - \sum_{j=1}^{\tau} \pi_j u_j - \beta\, \pi_{\tau+1} u_{\tau+1}}{1 - \sum_{j=1}^{\tau} \pi_j - \beta\, \pi_{\tau+1}},$$

或者等价地

$$u^* = \sum_{j=1}^{\tau} \pi_j u_j + \beta \pi_{\tau+1} u_{\tau+1} + \left(1 - \sum_{j=1}^{\tau} \pi_j - \beta \pi_{\tau+1}\right) u^*(\tau, \beta). \tag{8}$$

直观地说，假设发送者以概率 $1-\beta$ 和 β 分别在策略 s^{τ} 和策略 $s^{\tau+1}$ 之间进行混合。那么，如果发送者不披露额外信号时的收益为 $u^*(\tau,\ \beta)$，那么他们在策略 s^{nd} 与本段前面提到的混合策略之间无差异。我们用这个辅助值 $u^*(\tau,\ \beta)$ 来说明，在我们构造的均衡中，发送者不披露额外信号时的实际收益低于相应的 $u^*(\tau,\ \beta)$，其中 $\tau,\ \beta$ 由 $\sum_{j=1}^{\tau}\pi_j+\beta\pi_{\tau+1}=\hat{}q$ 确定，因此第二个激励约束得以满足。

根据定义，$u^*(\tau,0)=u_{\tau}^*$如（2）中所定义，且 $u^*(\tau,1)=u^*(\tau+1,0)$。回顾一下，$i$ 的选取满足 $u_i^*\equiv u^*(i,0)\geq u_{i+1}$。在引理5的证明中，我们证明了 $u_{\tau+1}^*\geq u_{\tau}^*$ 当且仅当 $u_{\tau}^*\geq u_{\tau+1}$。通过归纳论证可知，对每个 $\tau\geq i$，都有 $u_{\tau+1}^*\geq u_{\tau}^*$。[14] 根据定义，对每个 $\tau\geq i$，$u^*(\tau,\beta)$关于 β 递增，因为

$$\begin{aligned}\frac{\partial u^*(\tau,\beta)}{\partial\beta}=&\pi_{\tau+1}\frac{u^*-u_{\tau+1}+\sum_{j=1}^{\tau}\pi_j(u_{\tau+1}-u_j)}{(1-\sum_{j=1}^{\tau}\pi_j-\beta\pi_{\tau+1})^2}\\=&\pi_{\tau+1}\frac{(1-\sum_{j=1}^{i}\pi_j)u_i^*-\sum_{j=i+1}^{\tau}\pi_j u_j-\sum_{j=\tau+1}^{k}\pi_j u_{\tau+1}}{(1-\sum_{j=1}^{\tau}\pi_j-\beta\pi_{\tau+1})^2}\geq 0,\end{aligned}$$

其中，不等式由 $u_i^*\geq u_{i+1}$推出。对于单调披露规则，这意味着函数 $u^*(\tau,\ \beta)$在披露概率介于$q^*=\sum_{j=1}^{i}\pi_j$ 和交点$\hat{}q$之间时严格递增。

由于 $\mathbb{U}$ 随 q 递减，对于满足 $q>q^*$ 的每个 $(q,\ v)\in\mathbb{U}$，都有 $v\leq u_{i+1}$。因此，假设对于某个 $\tau\geq i$ 和 $(\hat{}q,\ \hat{}v)\in\mathbb{U}$，交点为 $\hat{}q=\sum_{j=1}^{\tau}\pi_{\tau}+\beta\pi_{\tau+1}$，那么 $\hat{}v\leq u_{i+1}\leq u^*(i,\ 0)\leq u^*(\tau,\ \beta)$，其中最后一个不等式根据 $u^*(\tau,\ \beta)$ 的单调性得出。这意味着，$\mathbb{U}$ 与 $\mathbb{V}$ 之间任何满足 $\hat{}q>q^*$ 的交点都将满足第二个激励约束。

[14]基础情形 $\tau=i$ 成立。假设 $u_{\tau+1}^*\geq u_{\tau}^*\geq u_{\tau+1}$，则 $u_{\tau+1}^*\geq u_{\tau}^*\geq u_{\tau+1}>u_{\tau+2}$，这意味着 $u_{\tau+2}^*\geq u_{\tau+1}^*$。

参考文献

Ali, S. N., A. Kleiner, and K. Zhang(2024)：《从设计到披露》，工作论文。Arieli, I. and C. Stewart (2025)：《没有承诺的贝叶斯劝说》，工作论文。Best, J. and D. Quigley (2024)：《长期说服》，《政治经济学期刊》，132，1740–1791。Corrao, R. and Y. Dai(2025)：《中介沟通的界限》，工作论文。DeVito, N. J., S. Bacon, and B. Goldacre(2020)：《遵守在 ClinicalTrials.gov 上报告临床试验结果的法律要求：一项队列研究》，《柳叶刀》，395，361–369。Dye, R. A.(1985)：《非专有信息的披露》，会计研究杂志，123–145。Dziuda, W.(2011)：《策略性论证》，《经济理论杂志》，146，1362–1397。Felgenhauer, M. and P. Loerke(2017)：《私人实验与贝叶斯劝说》，国际经济评论，58，829–856。Felgenhauer, M. and E. Schulte (2014)：《策略性私人实验》，《美国经济期刊：微观经济学》，6，74–105。Forges, F. (1986)：《沟通均衡的一种方法》，《计量经济学》，1375–1385。Fudenberg, D. and J. Tirole (1991)：《完美贝叶斯均衡与序贯均衡》，《经济理论杂志》，53，236–260。Gao, Y. (2025)：《从选择性披露的数据中进行推断》，工作论文。Goltsman, M., J. Hörner, G. Pavlov, and F. Squintani(2009)：《调解、仲裁与谈判》，《经济理论杂志》，144，1397–1420。Henry, E. and M. Ottaviani (2019)：《研究与审批流程：说服的组织》，《美国经济评论》，109，911–955。Kamenica, E. and M. Gentzkow (2011)：《贝叶斯劝说》，《美国经济评论》，101，2590–2615。Kamenica, E. and X. Lin(2024)：《沟通中的承诺与随机化》，工作论文。

Koessler, F. and V. Skreta (2023)：《知情信息设计》，政治经济学期刊，131，3186–3232。Kreps, D. and R. Wilson (1982)：《序贯均衡》，《计量经济学》，863–894。Kreutzkamp, S. and Y. Lou (2025)：《无事后承诺的说服》，《经济理论杂志》，106058。Lin, X. and C. Liu (2024)：《可信说服》，政治经济学期刊，132，2228–2273。Lipnowski, E. and D. Ravid (2020)：《动机透明的廉价交谈》，《计量经济学》，88，1631–1660。Lipnowski, E., D. Ravid, and D. Shishkin (2022)：《通过弱制度进行说服》，政治经济学期刊，130，2705–2730。——— (2025)：《完美贝叶斯说服》，《政治经济学杂志：微观经济学》，即将出版。Lou, Y.(2023)：《私人实验、数据截断与可验证披露》，arXiv预印本 arXiv:2305.04231。Lyu, Q. and W. Suen (2025)：《廉价交谈中的信息设计》，工作论文。Mathevet, L., D. Pearce, and E. Stacchetti (2024)：《声誉与信息设计》，工作论文。Myerson, R. B. (1982)：《广义委托代理问题中的最优协调机制》，《数理经济学杂志》，10，67–81。Pei, H. (2023)：《具有私人说谎成本的重复沟通》，《经济理论杂志》，210，105668。——— (2025)：《具有内生记录的社区执行》，《经济研究评论》，即将出版。Perez-Richet, E. (2014)：《中期贝叶斯劝说：初步探索》，《美国经济评论》，104，469–474。Rappoport, D. and V. Somma(2017)：《激励信息设计》，工作论文。Titova, M. and K. Zhang(2025)：《利用可验证信息进行说服》，《经济理论杂志》，106102。Zapechelnyuk, A. (2023)：《论未获信息与已获信息委托人的信息设计等价性》，《经济理论》，76，1051–1067。