arXiv:2303.01921v2 [physics.soc-ph] 4 Feb 2024

信任：独自与共同

BENEDIKT V. MEYLAHN、ARNOUD V. DEN BOER 和 MICHEL MANDJES

Abstract. 我们研究这样一个问题：某个代理人持续面临是否给予某机构信任的决策。代理人运用贝叶斯学习来估计该机构的真实可信度，并基于短视理性决定是否给予信任。借助随机游走理论中的要素，我们明确推导出此类代理人在关系中的某个时点停止给予信任的概率，以及以其停止给予信任为条件，给予信任所持续时间的期望值。随后，我们继续对两个 信任者进行建模，他们各自与该机构保持关系。我们考察两种自然的主体间沟通模型。在第一种（“可观察奖励”）中，主体彼此分享与该机构相关的经历；而在第二种（“可观察行动”）中，主体只是观察邻居的行动，也就是说，是否给予信任。在与单个代理人的情形相同的假设下，我们描述了两种不同沟通模型下主体信念的演变。尽管只有两个主体，停止给予信任的概率和系统中的期望时间仍无法得到显式表达。因此，我们开展仿真研究，以比较不同沟通方式对信任动态的影响。

我们发现，在这两种沟通模型中，一对主体比单个代理更有可能了解某个机构的真实可信度。主体之间的通信有助于与值得信赖的机构建立长期信任，也有助于及时退出与不值得信赖的机构之间的信任关系。与人们可能预期的相反，我们发现，较少的信息（观察彼此的行动，而非经历）有时对主体有利。

Keywords. 信任、学习、嵌入性、社会影响、制度信任

Affiliations. 阿姆斯特丹大学 Korteweg-de Vries 数学研究所；Science Park 904, 1098 XH 阿姆斯特丹；荷兰（联系方式：Benedikt V. Meylahn b.v.meylahn[at]uva.nl）。

Acknowledgments. 本研究得到欧盟地平线2020研究与创新计划 Marie Sk-lodowska-Curie 项目资助（拨款协议编号 945045），以及 NWO 引力项目 NETWORKS 资助（拨款编号 024.002.003）。

本项工作在 SURF 合作社的支持下，依托荷兰国家电子基础设施开展。版本：2024年2月6日。

作者声明不存在需要披露的利益冲突。

1. Introduction

本文关注个人决定是否信任某个机构的过程。COVID-19 大流行凸显了信任和不信任各国政府的重要性：某个群体对政府的不信任，导致该群体的个人疫苗接种率较低 [1]。同样，对科学的不信任与疫苗犹豫密切相关 [2]。COVID-19 信息疫情期间假新闻的盛行 [3] 表明，点对点沟通可能是普遍不信任的罪魁祸首。尽管这些结果是在 COVID-19 大流行和疫苗犹豫的背景下得出的，但人们预计，类似的动态也适用于许多其他社会现象。虽然个人与机构之间的个人经验

很重要，但与个人同伴的互动似乎也发挥着关键作用。因此，我们研究信任者彼此之间的点对点关系背景下，个体对机构的信任动态。

A 信任情境 是指个人必须承担一定的“风险”，以观察另一方的可信度的情境 [4, Ch. 1]。在某些情况下，这种风险涉及将资源交由他人支配。与机构的关系以信任为特征：在此类关系中，信托者通常会形成对该机构可信度的直觉。例如，我们可以考察个人与社交媒体公司之间的关系。个人因此将自己的数据（资源）交由公司掌管，作为回报，他们可以发布内容，并访问他人发布的帖子。个人观察到的可信度，取决于他们日常使用社交网络的经验，具体表现为（没有）听闻社交媒体公司将其信息出售给第三方广告公司。再举一个例子：考虑是否接种某种病毒疫苗的决策。对于那些相信科学共识会以他们的最大利益为重的个人来说，这项决策或许根本不需要信任。对于其他对科学机构和／或医疗卫生治理持怀疑态度的个人来说，在政府疫苗接种站接种疫苗，可能会被视为对自身健康乃至个人信息的一种风险。

这类例子促使我们将决定是否给予信任的问题表述为一个学习问题。代理人会了解其交互伙伴的可信度这一观点，也得到了社会学理论的支持。Buskens 和 Raub [5]撰写的章节很好地概述了将信任视为学习问题的观点，以及相关的实验性证据。此外，个人在面临是否给予信任的决策时，很少处于孤立状态。相反，人的行动嵌入于一个网络之中，而该网络很可能会影响其中个人的决策行为。因此，我们考察点对点沟通对信任问题动态的影响，这一点也得到了实证文献的支持（参见 [6, 7]）。特别是，Wang 和 Yu [8] 发现，无论是在经历得到交流时，还是在代理人仅仅观察他人行为时，这种影响都会发挥作用。我们将研究限定在两个代理人之内，因为这能够展现主体间通信所依托机制的本质。这样，我们便可以构建具有精确贝叶斯学习[1]的模型，而不是外生地对一个代理人向另一个代理人发送的信号所产生的影响施加模型假设；这里采用的是短视[2]理性这一基本假设。即使在两个代理人的情形下，这种方法也会产生精细而复杂的问题：代理人会解读邻居的行动，同时知道这些行动受到了他们自己迄今为止所采取行动的影响。由于这种相互影响，在双代理人设置中，我们模型中关键量的评估难以得到简单且显式的表达式。事实上，对这些量进行评估需要大量数值积分，因此我们在数值实验中使用了大型计算设施的一个集群。

1.1. 相关文献。 我们的研究与两个文献领域相关：社交网络学习和社交网络信任。下面我们将简要介绍相关文献，尽管这一介绍并不全面。与我们研究相关的工作涵盖多个研究学科，每个学科都有其自身的方法和一套

[1]生物感知领域的研究表明，人类的行为往往符合贝叶斯方式。关于其影响的讨论以及研究结果列表，请参阅 Knill and Pouget [9].[2]短视决策不仅是文献中的常见假设（参见 [10, 11, 12, 13, 14]），而且还有一些实验性证据表明，人类存在半短视行为 [15, 16]。

“合理假设”。相关领域包括但不限于经济学、理论社会学、社会物理学和机器学习。除了对这些动态的建模之外，关于网络影响信任的实证研究甚至延伸到了市场营销领域。

1.1.1. 社交网络学习。社交网络学习研究的是这样一类主体：他们通过选择行动来优化一个未知的目标函数，并通过私人信号和／或其他主体发出的信号进行学习。在各种建模决策下，研究关注的是一群主体学习最佳行动方案的速度，以及在理性假设下，这群主体是否可能采取次优行动（这往往是社会影响的结果）。建模决策涉及学习中的理性程度、主体之间的通信，以及主体的私人信号是否取决于其采取的行动。表1概述了相关论文的建模决策 [12, 17, 14, 18], ，这些论文均在社交网络学习框架下考察主体群体。倒数第二列“通信”指主体彼此分享的内容，可能是每轮收到的私人学习信号、信念分布，或仅仅是他们采取的行动。最后一列“信号”指主体收到私人信号这一过程，是否独立于其采取的行动，或取决于这些行动。

论文	理性	通信	信号
Bala 和 Goyal [12]	短视	条件行动和信号	
Molavi等 [17]	不完美记忆信念		独立
Harel等 [14]	短视	行动	独立的
黄等人 [18]	非短视的	行动	独立的

TABLE 1. 精选社交网络学习文献中的建模决策。

Bala 和 Goyal 的论文中的主体 [12] 是短视的；此外，他们不会根据其直接邻居所采取的行动，对这些邻居的邻居的结果作出任何推断。作者发现，在连通网络中，主体的行为渐近收敛。Molavi 等人 [17], 、Harel 等人 [14] 以及 Huang 等人 [18] 均考虑了这样一些模型：主体接收到的私人信号与其所采取的行动相互独立。在此背景下，Molavi 等人 [17] 说明了贝叶斯学习背景下的不完全记忆如何与 DeGroot 的开创性模型相关联 [19]。在完全连通的人群模型中，Harel 等人 [14] 和 Huang 等人 [18] 将通信仅限于行动的模型的学习速率，与主体之间可就信号和信念进行开放通信的参照模型联系起来。

脱离网络环境后，Correa 等人 [20] 以及Fu和Le Riche [21] 考察了一种模型：来自某个人群的依次到达的智能体根据基于前面智能体观察结果形成的关于产品质量的贝叶斯信念，选择是否购买产品。Correa 等人 [20] 计算了不完全学习发生的概率，即智能体错误地认为产品质量较低的概率。 [20] 与我们的设定之间的一个重要区别在于， [20] 中的智能体知道产品质量来自两个已知值的集合。这意味着，在 [20] 中，智能体只需识别质量取两个值中的哪一个；而在本文中，相关参数（下文称为可信度）可以取单位区间内的任意值。Fu和Le Riche [21] 将类似的学习问题纳入一个内生增长市场模型，其中智能体将质量未知的新产品与质量已知的旧产品进行比较，由此得到的结果是：存在一些可实现的均衡，在这些均衡中，产品的真实质量始终未知。

此外，还有一批文献专门研究确保社会学习的条件。在这些论文中，主体通常只接收一个信号，并依次采取行动，而这些行动会被其余主体中的一部分或全部观察到。在Banerjee和Fudenberg [22],的情形下，观察结构采取代表性样本的形式，而人群是一个连续统。Acemoglu 等人 [23] 引入网络结构，以确定主体能够观察到哪些行动。关于主体由全连接图表示的模型综述，感兴趣的读者可参阅Birkchandani 等人 [24]。

据我们所知，社交网络学习文献尚未研究过以下相关设定：在一个模型中，主体采取短视理性的行动，观察邻居的行动或信号，并拥有取决于所采取行动的私人信号。这类模型也与社交网络信任文献相关，因为信任关系要求信号与行动之间存在依赖关系。

1.1.2. 社交网络信任。 社交网络信任通常关注一对信任者和受托者参与迭代信任博弈时的行为。研究重点在于信任破裂的可能性，以及描述信任者和受托者的理性策略。与我们的研究尤其相关的是这样一些研究：人群中既有始终信守信任的值得信任的受托者，也有采取策略性行动的机会主义受托者。在这类模型中，让主体学习值得信任的受托者所占比例，是一种自然的扩展。

这一文献脉络中的重要范例见 Bower 等人 [25],Buskens [26], 和 Frey 等人 [27]。Bower等人 [25] 描述了一系列两轮信任博弈的均衡策略：每次都会配对一对新的信任者和受托者，进行两轮信任博弈。除了学习值得信任的受托者的真实比例外，在同一场配对中，信任者和受托者还会从第一轮学习到第二轮。Buskens [26] 通过考察一对都与同一受托者有关系的信任者，扩展了这一分析。主要发现是，只有当两个主体分享信息的概率都很高时，信任者的信任投入和受托者的信守行为（在均衡策略中）才会随着信任者之间分享信息的概率增加而增加。Buskens 等人 [28]的实验研究证实了这些发现。Frey 等人 [27] 进一步拓展了这项理论研究，假设信任者之间的连接需要付出成本。作者确定了在均衡策略中哪些博弈参数会促成或排除对此类连接的投资。

与这类文献不同，我们考虑的是一个受托人主体，例如某个机构，其行为并非以策略方式建模，并且会与多个信任主体互动。按照例如 [26, 27, 28], 一个受托人与两个信任者互动的情形，是一个值得考虑的自然场景。

1.1.3. 相关性与视角。 本节将详细阐述这两类不同文献流派对于我们所提出模型的相关性。在主题上，我们立足于社交网络信任文献；在方法论上，我们的方法与社交网络学习中所采用的方法相似。我们首先介绍这两类文献流派中的观点，随后说明我们的工作如何融入这一领域。

社交网络信任文献详细描述了策略性信任者与两类受托者之一之间的互动，例如始终恪守信任的“友好型”受托者，以及试图“欺骗”信托者、以便在某个时刻滥用信任的“策略型”受托者（参见 [25, 26, 27]）。 “友好型”受托者的比例可能是已知的（如 [26, 27]所述），也可能是未知的（如 [25]所述）。

信任者正在学习他们所交互的受托者属于何种类型，并且有时也会学习“友好型”受托者所占的比例。一旦某位受托者在点对点情境中滥用过一次信任，他们就会暴露出自己不值得信任。该文献中一个常见的例子是，人们从一位关系疏远的熟人处或通过在线点对点服务购买二手车（参见Buskens [4,第 6]章），或者在未签订合同的情况下雇用一名非正式的看房人；在这两种情境中，信托者和受托人都在某种意义上是同伴。

在社交网络学习文献中，存在一种环境，它为主体可能采取的行动设定了顺序，并发出信号。主体采取行动并接收信号。他们利用信号更新对环境的信念，以期采取最优行动。这之所以属于社会学习，是因为主体会彼此交流各自的行动和/或信号。研究者则关注群体学习到最优行动的概率（参见 [12, 20, 21]），以及在不同通信形式下收敛到这一最优行动的速度（参见 [14, 18]）。社交网络学习的研究结果适用于主体持续更新信念并试图采取最优行动的情形。市场预测（并设定相应价格以实现利润最大化）、观点采纳（试图融入群体）以及信息传播（及时了解最新信息）都是这类情形的例子。

我们关注的问题是对机构的信任，以及点对点沟通对其产生的影响。因此，社交网络信任文献为我们的研究提供了主题背景。我们感兴趣的是信任丧失这一事件，主要关注信任者与机构互动的“非对称”情形。在这种情境下，将受托人的策略行为纳入考量并不那么自然：他们并未主动参与互动，而是“被动地”提供服务。我们不将受托人的行为建模为策略行为，而是让他们简单地从某个分布中抽取一个行动。

我们考察了社交网络学习文献，因为我们的模型与其中的模型在技术上相似。我们的目标是寻找适用于我们所研究情境的工作（即使采用了不同的术语）。在此过程中，我们发现自己的模型也填补了社交网络学习文献中的一个空白。此外，这也让我们更坚定地决定比较不同沟通形式对结果的影响。在社交网络信任文献中，通常假设主体总是完整地交流自己的经历或信念分布。不过，有人可能会认为，交流经历未必能提供恰当的视角：行动通常很容易观察到，而内在信念或个人经验则未必总会分享。这促使我们对两者进行比较。

由此形成的建模框架自然适用于我们所关注的情境：个人在点对点影响的背景下信任机构。

1.2. 贡献。 我们的模型旨在弥补上述文献中的空白。为了对个人与机构之间信任问题的动态中，点对点沟通所产生的影响进行建模，我们借鉴了这两个领域的文献。按照社交网络信任文献的做法，学习信号取决于所采取的行动。我们通过采用短视理性决策和贝叶斯学习程序，遵循社交网络学习文献。我们研究了这两个相关领域的文献中出现的两种主体间沟通模型。在第一种模型（“可观察奖励”）中，主体彼此透露自己与机构打交道的经验；而在第二种模型（“可观察的行动”）中，主体只观察邻居是否给予信任，即是否信任机构。在这两种沟通模型中，我们都描述了如何理性使用信息来更新信念。理性的程度起到基准作用

而非对实际人类行为的描述。相关计算很快就会变得复杂，但能够为信念更新过程中如何完全理性地使用信息提供有益的参考。

我们不对受托人的动机作任何假设；受托人只是以概率 $\theta \in (0, 1)$ 诚实行事，从而将情形推广到该概率取两个值之一的情况，如 [20] 所示。我们关注这样一种设定：信任主体与机构进行交互，而这些机构的行为无法像 [25, 26, 27, 29] 中那样，在个人交互层面通过策略模型来刻画。我们还考虑 [26, 27]，中通信方式的一个更一般版本，在这种版本中，主体可以获取邻居的行动以及由此产生的结果。

我们描述信息的使用和短视决策，而不依赖于 [17, 14, 18] 中所示的、独立于所采取行动的信号。行动的不对称性是信任问题的自然特征，这意味着可能出现不完全学习。因此，我们关注趋于真相的概率，而不仅仅关注其趋近速度。

在我们的研究中，我们运用随机游走理论领域的技术分析单代理情形的动态。然而，这类分析技术无法推广到双智能体模型，因此我们依靠蒙特卡罗模拟来分析这些模型。由于代理人解读彼此行动所需的数值积分在个人计算机上会耗费难以承受的计算时间，我们不得不使用计算设施 SURFsara 的 Lisa 集群。

我们发现，彼此之间进行通信的两个代理人往往能更快做出“正确”决策。换言之，通常来说，关系有助于代理人的样本路径所带来的益处，超过了因通信而导致一方代理人“运气不好”便意味着双方都“运气不好”的样本路径所带来的损失。

我们的实验表明，可观察奖励模型并不总是比可观察行动模型“更好”。哪种模式对代理人最有帮助，取决于人们关注的是长期做出正确决策，还是尽可能迅速地确保与不值得信赖的机构结束关系。此外，我们还找到了一个参数设定，在该设定下，可观察行动模型中的退出概率低于 可观察奖励模型。这意味着，与人们可能预想的相反，较少的信息有时可能对代理人更有利，而且这两个模型之间并不存在单调的优劣排序。

1.3. 论文结构。 在 §2 中，我们描述模型的基本要素及相关解释。随后，在 §3 中，我们进一步描述模型，并分析单个代理的情形（证明见附录 A）。在此，我们特别关注那些能够对这种信任关系终止的概率进行解析分析的子情形。在 §4 中，我们为两个代理与同一机构建立信任关系的情形，提出两种模型：可观察奖励和可观察的行动。在 §5 中，我们讨论用于研究双代理模型的实验设置，并在 §6 中讨论实验结果。最后，我们在 §7 中讨论这些结果的影响及其与现有文献的关联。

2. Model for a single agent

本节介绍单个代理模型，它是后续各节所讨论的双代理模型的核心。我们考虑这样一种情形：代理人有多次机会信任某个机构。机构的行为由单个参数 $\vartheta \in [0, 1]$建模，即真实的可信度，定义为信任得到兑现的概率

（因此，其补数 $1-\vartheta$ 就是信任遭到滥用的概率）。如果没有给予信任，机构便无须采取行动。可以将这种行为理解为机构兑现信任的效能，并隐含地假定机构在每一轮中都试图这样做。需要指出的是，我们承认对机构的刻画具有很高的抽象程度，而我们主要关注的是代理人在这种情形下的行为。在每一轮 $t \in \mathbb{N}$ 中，代理人从行动集 $\mathcal{A}=\{0,1\}$中选择一个行动，其中 $A_t=1$ 表示代理人在第 t 轮给予信任，而 $A_t=0$ 表示代理人退出信任关系。我们定义随机变量 $X_t, \forall t \in \mathbb{N}$，表示机构在第 t 轮的行动是（或将会是）兑现还是滥用可能已给予或未给予的信任：

$$X_t=\begin{cases}1, & \text{with probability } \vartheta\\ 0, & \text{with probability } 1-\vartheta.\end{cases} \tag{1}$$

随机变量 $\{X_t: t\in\mathbb{N}\}$ 是独立同分布的，尤其与代理人的行动相互独立。重要的是，代理人仅在 $A_t=1$的轮次中观察到 X_t。如果信任得到履行，代理人获得效用 r （奖励）；而如果信任遭到滥用，效用增益为 $-c$ （成本）。我们假设 c 和 r 是正整数；请注意，如果 $r,\ c\in\mathbb{Q}$ 我们只需将二者同时乘以其分母的乘积，即可得到整数。代理人给予机构信任时的效用为 $rX_t\mathbf{1}_{\{A_t=1\}}-c\,(1-X_t\mathbf{1}_{\{A_t=1\}})$，因此期望效用为 $r\vartheta-c\,(1-\vartheta)$。正如学习文献中广泛采用的做法（例如，Sebenius 和 Geanakoplos [10],Parikh 和 Krasicki [11], Bala 和 Goyal [12], Keppo 等人 [13] 以及 Harel 等人。[14]），我们让代理人采取短视理性。这意味着在每一轮 t 中，他们只考虑立即行动的期望效用，而不考虑在 $t+1$、$t+2$.....轮中采取行动可能带来的效用，也不考虑通过采取行动 $A_t=1$所获得的信息的效用。当代理人相信该效用的期望值非负时，他们就会给予信任。此外，代理人以参数为 α、 $\beta\in\mathbb{N}$的贝塔分布先验信念 P_0 开始，因此他们最初相信 ϑ 的概率密度由下式给出

$$P_0(\theta)=B(\theta;\alpha,\beta):=\frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\mathrm{d}y},\quad \theta\in[0,1]. \tag{2}$$

$\widehat{\ }\ \theta$ 的期望的初始估计值为 $\vartheta_0=\mathbb{E}[B(\alpha,\beta)]=\alpha/(\alpha+\beta)$。随着更多信息的出现（在轮次 $t>0$ 中给予信任，并且随后信任得到兑现或遭到滥用），代理人会以贝叶斯方式更新这一信念分布。令

$$\widehat{S}_t=\sum_{s=1}^{t}X_s\mathbf{1}_{\{A_s=1\}}, \tag{3}$$

表示截至时间 t，代理人观察到信任得到兑现的次数。同样，令

$$\widehat{F}_t=\sum_{s=1}^{t}(1-X_s)\mathbf{1}_{\{A_s=1\}}, \tag{4}$$

表示截至时间 t，代理人观察到信任遭到滥用的次数。然后，应用贝叶斯法则，求得代理人在第 t 轮结束时持有的信念分布，从而得到

$$P_t(\theta)=\frac{\theta^{\widehat{S}_t}(1-\theta)^{\widehat{F}_t}P_0(\theta)}{\int_0^1 y^{\widehat{S}_t}(1-y)^{\widehat{F}_t}P_0(y)\mathrm{d}y},\quad \theta\in[0,1]. \tag{5}$$

将时间 t 的可信度估计值记为

$$\widehat{\vartheta}_t = \mathbb{E}_{\theta\sim P_t(\theta)}[\theta] = \frac{\alpha + \widehat{S}_t}{\alpha + \beta + \widehat{S}_t + \widehat{F}_t}, \quad \forall t \in \mathbb{N}, \tag{6}$$

那么对于 $t = 1$，2，.....我们有：

$$A_t = \begin{cases} 1, & \text{if } r\widehat{\vartheta}_n - c(1-\widehat{\vartheta}_n) \geq 0, \quad \forall n \in \{0, 1, \ldots, t-1\}, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

2.1. 关注的量。我们关注代理人的退出，即他们停止给予信任的情况：$\{A_t = 0\}$。随机变量 τ 表示首次采取“不再给予信任”这一行动之前给予信任的轮次数：

$$\tau := \inf\{t \in \mathbb{N} \cup \{\infty\} : r\widehat{\vartheta}_t - c(1-\widehat{\vartheta}_t) < 0\}. \tag{8}$$

根据 (8) 中 τ 和 (7) 中 A_t 的定义，我们有

$$A_t = 1, \quad \forall t \leq \tau \quad \text{and} \quad A_t = 0, \quad \forall t > \tau. \tag{9}$$

因此我们注意到，对于所有 $t > \tau$，$\widehat{\vartheta}_t = \widehat{\vartheta}_\tau$ 都成立。这一结果从模型动力学来看十分自然：代理人一旦采取不再给予信任的行动，就无法观察该行动的结果，因此也不会调整其估计值。这将代理人的行动与估计值联系起来，反之亦然。一个机构可能具有真实的可信度ϑ，使得 $r\vartheta - c(1-\vartheta) > 0$，这意味着，了解 ϑ 真实值的代理人会永远给予信任。在这种情况下，代理人的估计值 $\widehat{\vartheta}$ 也可能持续满足条件 (7)，这意味着他们永远不会失去信任。在这种情况下，退出并非必然，因此一个特别值得研究的概率是事件 $\{\tau < \infty\}$ 的概率。我们将其记为

$$p_{\text{quit}} := \mathbb{P}(\tau < \infty). \tag{10}$$

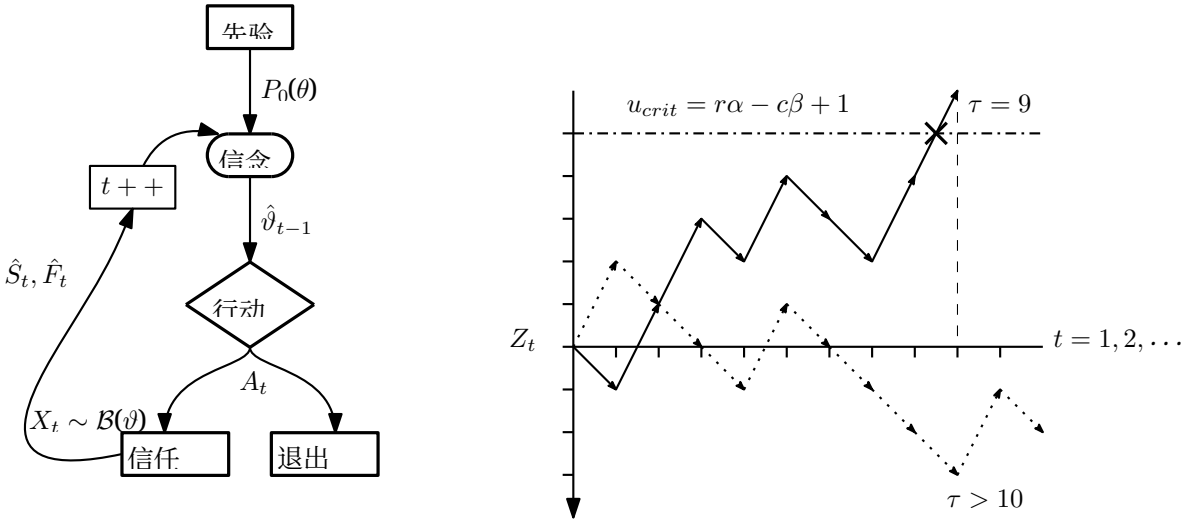
此外，我们还关注在以退出为条件的情况下，退出前在系统中花费的期望时间：

$$q := \mathbb{E}[\tau \,|\, \tau < \infty]. \tag{11}$$

现在，我们参照图1a中的图示，回顾一下动态过程。在时间 t 结束时，决定代理人信念分布的两个要素是其先验 P_0，以及包含在随机变量 S_t 和 F_t中的交互历史。这些共同构成对机构 ϑ_t 可信度的估计，代理人据此选择在第 $t+1$轮次信任该机构，或退出信任关系。如果代理人选择信任（从而继续这一过程），机构将作出回应：要么履行信任（概率为 ϑ），要么辜负信任（概率为 $1-\vartheta$）。请注意，模型中的所有随机性都源自机构的回应。

3. Trusting alone

本节分析 §2 中描述的单智能体模型的动态。我们关注退出概率 p_{quit} 和预期退出时间 q（以退出发生为条件）。我们考察单个代理人定期选择是否信任机构时所呈现的动态。本节从 §3.1 开始，将模型重新解释为具有吸收屏障、步长可能不对称且步概率也可能不对称的随机游走。随后在 §3.2 中，我们分析该随机游走模型的若干特殊情况，重点评估指标 p_{quit} 和 q。



(A) 单智能体模型。 (B) 示例1中的样本路径。

FIGURE 1. (A) 单智能体学习者模型的概念示意图，以及 (B) 在随机游走解释下模型动态的两条示例路径。 $\mathcal{B}(\vartheta)$ 表示参数为 ϑ 的伯努利分布。

3.1. 模型重新解释。 我们将 §2 中描述的模型解释为一维随机游走，其步长为 $-r$ 和 $+c$，对应的概率分别为 $\vartheta \in (0,1)$ 和 $1-\vartheta$。对于持有初始贝塔信念分布、参数为 α 和 β 的代理人，决策标准变为

$$r\left(\frac{\alpha+\widehat{S}_t}{\alpha+\beta+\widehat{S}_t+\widehat{F}_t}\right) - c\left(1-\frac{\alpha+\widehat{S}_t}{\alpha+\beta+\widehat{S}_t+\widehat{F}_t}\right) \geq 0. \tag{12}$$

经过变换，我们发现，要建立信任，代理人需要

$$c\widehat{F}_t - r\widehat{S}_t \leq r\alpha - c\beta, \tag{13}$$

我们将其解释为一个随机游走

$$Z_t := c\widehat{F}_t - r\widehat{S}_t, \quad \text{for } t = 0, 1, \ldots, \tag{14}$$

从 $Z_0 = 0$ 开始，并在“临界”位置设置一个吸收屏障

$$u_{\text{crit}}(\alpha, \beta, c, r) := r\alpha - c\beta + 1. \tag{15}$$

请注意，只要 u_{crit} 的值相同， α 和 β 的不同取值在 p_{quit} 和 q_{quit} 方面对应的行为也相同。(15) 中多出的一个单位，是因为建立信任的决策标准包含等号（预期奖励必须大于或等于零），而吸收屏障的表述描述的是第一个不建立信任的位置。请注意，先验的所有影响都包含在 u_{crit} 中。这也使我们能够将退出概率重新表述为

$$p_{\text{quit}} \equiv p_{\text{quit}}(\alpha, \beta, c, r, \vartheta) = \mathbb{P}(\exists t : Z_t \geq u_{\text{crit}}), \tag{16}$$

随机游走到达吸收屏障的概率。为便于参考$\widehat{\ }$，我们将 $c/(c+r)$ 称为 θ_{crit}，因为与 $Z_t \geq u_{\text{crit}}$ 等价的条件是 $\vartheta < \theta_{\text{crit}}$。作为示例，我们在图1b中给出了这种随机游走可能采取的两条样本路径。

示例1（随机游走解释）。 考虑由参数值 $c = 2$、 $r = 1$、 $\alpha = 8$ 和 $\beta = 2$产生、在 $u_{\text{crit}} = 5$ 处设有吸收屏障的情形。我们将其解释为具有临界可信度 $\theta_{\text{crit}} = 2/3 = 0.67$。第一条路径（以实线箭头表示）在 $t = 9$处到达吸收屏障，而第二条路径（以虚线箭头表示）在所示时间步内没有到达吸收屏障，因此具有 $\tau > 10$。 S_t 和 Z_t 的$\widehat{}$值以及相应的可信度估计值$\widehat{\vartheta}_t$列于表2中。

♣

TABLE 2. 作为随机游走的样本路径和代理信念。

(A) 图1b中的实线。

t	0	1	2	3	4	5	6	7	8	9
$\widehat{S}_t$	0	1	1	1	2	2	3	4	4	4
Z_t	0	-1	1	3	2	4	3	2	4	6
$\widehat{\vartheta}_t$	0.8	0.82	0.75	0.69	0.71	0.67	0.69	0.71	0.67	0.63

(B) 图1b中的虚线。

t	0	1	2	3	4	5	6	7	8	9	10	11
$\widehat{S}_t$	0	0	1	2	3	3	4	5	6	7	7	8
Z_t	0	2	1	0	-1	1	0	-1	-2	-3	-1	-2
$\widehat{\vartheta}_t$	0.8	0.73	0.75	0.77	0.79	0.73	0.75	0.76	0.78	0.79	0.75	0.76

关于退出概率 p_{quit}，我们首先考虑 $\vartheta < c/(c+r)$ 的情况。请注意，在这种情况下，给予信任的实际期望效用为负，并且过程Z_t 向吸收屏障漂移，这意味着最终以概率 1 到达该屏障。

引理1（保证退出）。 如果 $\vartheta < c/(c + r)$，则 $p_{\text{quit}} = 1$， 并且 $\tau < \infty$.

证明（见附录 A）高度依赖大数定律。本节其余部分的研究聚焦于（更有趣的）情形 $\vartheta > \theta_{\text{crit}}$。这一情形尤为重要，因为它代表了代理人的最优行动是无限期地给予信任，但他们未必会这样做的情况。我们研究具有一般吸收屏障 $u \geq 0$ 的随机游走，以找出不同 u 值下 p_{quit} 之间的递推关系。针对相关参数值，我们利用适当的 u，求出退出信任关系的概率以及退出发生的时间。

根据迄今描述的模型，代理人要么继续无限期地给予信任，并了解可信度的真实值；要么在某个时刻 $t < \infty$ 退出。

引理 2 (收敛或退出)。 参与上述信任关系的单个代理，在 $\vartheta > \theta_{crit}$ 下要么 （A） 在某个时刻 $\tau < \infty$ 退出，即

$$A_t = 1, \quad \forall t < \tau, \quad and \quad A_t = 0, \quad \forall t \geq \tau, \tag{17}$$

或（B）继续无限期地给予信任，并使他们的估计值趋近于真实的可信度，即，

$$A_t = 1, \quad \forall t = 1, 2, \ldots, \quad and \quad \lim_{t\to\infty} \widehat{\vartheta}_t = \vartheta. \tag{18}$$

这意味着

$$\mathbb{P}(\widehat{\vartheta} \not\to \vartheta \ \& \ \tau = \infty) = 0. \tag{19}$$

证明由估计可信度值的定义和大数定律得出，见附录 A。

3.2. 结果。 在本小节中，我们研究代理人退出信任关系的概率。我们通过求出随机游走在某个水平 u 处的吸收概率来实现这一目标。我们将 $\pi(u)$ 定义为：从起点 $Z_t = 0$ 到位于 $u \geq 0$ 的吸收屏障的距离为 u 时，触及该吸收屏障的概率：

$$\pi(u) \equiv \pi(u, c, r, \vartheta) := \mathbb{P}(\exists t : Z_t \geq u), \quad \forall u \in \mathbb{N}_0. \tag{20}$$

为便于阅读，我们省略对 c，r 和 ϑ 的依赖。请注意，u 可以等于 0，这对应于必然发生吸收的情形。关于 $\pi(u)$ 的分析根据 r 和 c 分为三种情况。其中两种情况可通过解析方法刻画，第三种情况则采用数值近似。使用 $u = u_{\text{crit}}$ 可得退出概率，即也就是，

$$p_{\text{quit}}(\alpha, \beta, c, r, \vartheta) = \pi(u_{\text{crit}}, c, r, \vartheta). \tag{21}$$

在 §§3.2.1-3.2.2 中，我们对上述两种情况进行了分析：1）$r \in \mathbb{N}$ 和 $c = 1$，以及 2）$r = 1$ 和 $c \in \mathbb{N}$。附录 B 给出了针对 r，$c \in \mathbb{N}$ 这一情况的近似方案。之所以这样划分，是因为一般情况下无法求得闭式结果。此外，所介绍的两种情况所用的技术也大不相同，因此应分别看待。

3.2.1. 情形 $c = 1$，且 $r \in \mathbb{N}$。 在这种情形下，随机游走具有一种有用的无记忆性。观察可知，随机游走向上时每次只能上升一个层级，而向下时则可能跳过若干层级。为此，我们先定义随机游走在某个 $t \in \mathbb{N}$ 的情况下上升一个层级的概率：

$$\varrho := \mathbb{P}(\exists t : Z_t \geq 1). \tag{22}$$

根据强马尔可夫性质，我们注意到，达到 1 个层级之后的动态与此前的历史无关。这意味着，向上走 u 个层级的概率，恰好等于连续 u 次向上走 1 个层级的概率，因此$\pi(u) = \varrho^u$。退出概率的形式化表达式（包括对 ϱ 的刻画）将在下面的引理中给出，证明见附录 A。

引理 3（当 $c = 1$，且 $r \in \mathbb{N}$ 时的退出概率）。设 $\vartheta \geq 1/(1 + r)$。参数值为 $c = 1$、$r \in \mathbb{N}$ 的相应随机游走到达位于 u 的吸收屏障的概率满足

$$\pi(u, 1, r, \vartheta) = \varrho(\vartheta)^u, \quad \forall u \in \mathbb{N}_0, \tag{23}$$

其中，$\varrho(\vartheta) = \pi(1, 1, r, \vartheta)$ 是范围 $[0, 1)$ 内方程的唯一解

$$\varrho(\vartheta) = (1 - \vartheta) + \vartheta\varrho(\vartheta)^{r+1}. \tag{24}$$

要计算退出概率，我们使用 $u_{\text{crit}} = r\alpha - \beta + 1$ *in*

$$p_{\text{quit}}(\alpha, \beta, 1, r, \vartheta) = \pi(u_{\text{crit}}, 1, r, \vartheta). \tag{25}$$

在图2中，我们绘制了退出概率的数值结果，记为 $p_{\text{quit}}(\alpha, \beta, c, r, \vartheta) = \pi(u_{\text{crit}}, c, r, \vartheta)$，对应于子情形 $r = 1$、2、3。贝塔先验信念分布由形状参数 $\alpha = \beta = 2$ 给出。这些线条是通过引理 3 得到的理论结果，而点（带置信区间）则对应于模拟结果，并与分析结果相吻合。退出概率的形状可以这样解释：较低的可信度 ϑ 始终会导致更高的退出概率。引理 1 和引理 3 的结果表明，当 $\vartheta < \theta_{\text{crit}}$ 时，退出概率为 1，此后遵循（25）。

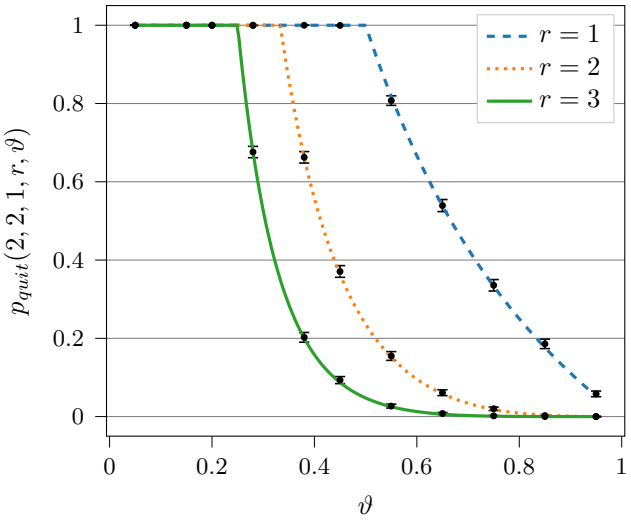


FIGURE 2. 单个代理退出的概率 p_{quit}（α，β，c，r，ϑ）随 ϑ 绘制，其中 r 取不同值，而 $\alpha = \beta = 2$和 $c = 1$。解析结果以线条表示，模拟结果（4 000 次迭代）以点表示，并附有95%置信区间。

现在我们将注意力转向代理人在系统中所花费的期望时间。注意，在随机游走表示法中，退出时刻 τ 可表示为

$$\tau(u) \equiv \tau(u, c, r, \vartheta) = \inf\{t \in \mathbb{N}_0 : Z_t \geq u\}. \tag{26}$$

于是，我们关注在 $\mathbf{1}\{\tau(u) < \infty\}$ 这一条件下 $\tau(u)$ 的期望；现在我们知道，该事件以 $\pi(u)$ 的概率发生。具体而言，我们研究以下量

$$\mathbb{E}[\tau(u) \,|\, \tau(u) < \infty] = \frac{\mathbb{E}\left[\tau(u)\mathbf{1}\{\tau(u) < \infty\}\right]}{\pi(u)}, \text{ for } u \in \mathbb{N}_0. \tag{27}$$

我们采用概率生成函数的技术来描述相关分布的期望。因此，我们考察一个与达到第一个层级的期望时间 $\tau(1)$相关的量：

$$\varphi(z) := \mathbb{E}[z^{\tau(1)}\mathbf{1}\{\tau(1) < \infty\}], \quad z \in [-1, 1], \tag{28}$$

并注意到

$$\mathbb{E}[z^{\tau(u)}\mathbf{1}\{\tau(u) < \infty\}] = \varphi(z)^u, \quad \forall u \in \mathbb{N}_0. \tag{29}$$

这（再次）源于一个事实：必须先通过所有 $\{1, \ldots, u-1\}$ 级别，才能到达 u。因此，到达 u^{th} 级别的预期时间，只需用到达第一级的期望时间乘以 u。

引理 4（当 $c = 1$ 且 $r \in \mathbb{N}$ 时的预期退出时间）。若 $r \in \mathbb{N}$，$c = 1$ 且$\vartheta \geq 1/(r+1)$ 则对所有 $u \in \mathbb{N}$，$\tau(u)$ 满足

$$\mathbb{E}[\tau(u) \,|\, \tau(u) < \infty] = u\frac{\varphi'(1)}{\varrho}, \quad \forall u \in \mathbb{N}, \tag{30}$$

其中 $\varphi'(z)$ 是函数 $\varphi(z)$ 的左导数。对于任意给定的 $z \leq |1|$，该函数是下列方程的唯一解：

$$\varphi(z) = z(1-\vartheta) + z\vartheta\,\varphi(z)^{r+1}, \tag{31}$$

其中 ϱ 是方程（*24*）的唯一解。特别地，

$$\mathbb{E}[\tau(u) \,|\, \tau(u) < \infty] = \frac{u\varrho}{(1-\vartheta)(r+1) - r\varrho}, \quad \forall u \in \mathbb{N}. \tag{32}$$

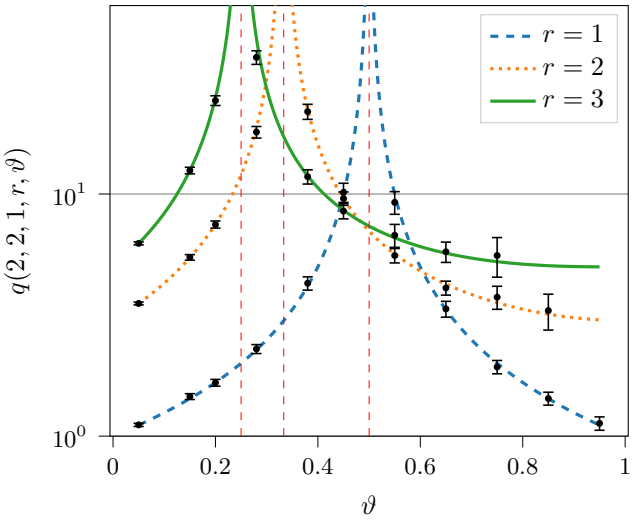


FIGURE 3. 在给定代理人在某个 $t < \infty$ 时退出的条件下，系统中的期望时间，针对 r 的不同取值绘制于 ϑ 上，同时 $\alpha = \beta = 2$和 $c = 1$ ，采用对数—线性坐标轴。解析结果以曲线表示，模拟结果以带有95%置信区间的点表示。

为求得预期退出时间（以退出为条件），我们在

$$q(\alpha, \beta, 1, r, \vartheta) = \tau(u_{\text{crit}}, 1, r, \vartheta). \tag{33}$$

要找到系统中的期望时间，首先需要针对 r 的适当取值，求解（24）中 $\varrho \in (0,1)$的唯一解。随后将该解与$u = u_{\text{crit}}$一并代入（32），从而得到所需的 $\tau(u_{\text{crit}}, 1, r, \vartheta)$表达式。详细内容请参见附录 A 中的证明。

我们在图 3 中绘制了代理人最终退出 $q(\alpha, \beta, 1, r, \vartheta)$时系统中的期望时间的数值结果和模拟结果，针对子情形 $r = 1, 2, 3$，采用对数—线性坐标轴。贝塔先验信念分布由形状参数 $\alpha = \beta = 2$刻画。我们注意到，期望时间逐渐增大并趋近于位于 $\vartheta = 1/(1 + r) = \theta_{\text{crit}}$上的渐近线。该渐近线的位置源于模型的性质：对于 $\vartheta = \theta_{\text{crit}}$ ，退出仍然是必然发生的，但所需时间的期望为无穷大（参见对称伯努利随机游走）。这种初始增长是合理的，因为随着 ϑ 增加，偏离退出的绕行概率也随之增大，尽管在该区域内代理人不可能永远逃脱。随后随着 $\vartheta > \theta_{\text{crit}}$ 的增加而下降，这是因为随着 ϑ 增加，一旦发生退出，退出就会更早发生。这是因为在关系中停留的时间越长，代理人接触到的无偏信息越多，而这些信息很可能表明该机构值得信任。

3.2.2. 情形 $r = 1$，以及 $c \in \mathbb{N}$。在这种情形下，我们注意到随机游走 Z_t 不再具有无记忆性：现在向上的步长为 $c \in \mathbb{N}$ ，因此沿途可能会跳过某些层级。这意味着，类型为 $\pi(u) = \varrho^u$ 的关系不再成立。因此，我们可以看出，无法保证在退出之前会经过哪些层级，所以不能采用我们在 $c = 1$ 情形中所依赖的方法。尽管如此，我们仍能得到以下结果。在这种情形下， u 的适当取值为 $u_{\text{crit}} = \alpha - c\beta + 1$。

引理 5 (当 $r = 1$，且 $c \in \mathbb{N}$ 时的退出概率)。若 $r = 1$、 $c \in \mathbb{N}$ 和 $\vartheta \geq c/(c+1)$，则退出概率为

$$p_{\text{quit}} = \pi(u_{\text{crit}}, c, 1, \vartheta), \tag{34}$$

其中 $\pi(u、\ c、\ 1、\ \vartheta)$ 满足

$$\pi(u) = \frac{\xi^{(u)}(0)}{u!}, \quad \forall u \in \mathbb{N}; \tag{35}$$

这里， $\xi^{(u)}(0)$ 是函数 $\xi(w)$ 关于 $w = 0$ 的 u^{th} 导数，其中 $\xi(w)$ 定义为

$$\xi(w) = \frac{[(1-\vartheta)(c+1)-1]w^2 + [2-(1-\vartheta)(c+1)]w - (1-\vartheta)w^{c+1} - \vartheta}{(1-\vartheta)w^{c+2} - (1-\vartheta)w^{c+1} - w^2 + (1+\vartheta)w - \vartheta}, \ \forall w \in (-1,1). \tag{36}$$

证明（见附录 A）利用了一种转换：将我们的随机游走的最大值与另一个带有 $c = 1$ 和 $r \in \mathbb{N}$ 的随机游走的最小值联系起来。

我们注意到， $\xi(u)$ 的形式为 $P(w)/Q(w)$，其中 $P(w)$ 和 $Q(w)$ 分别是次数为 $c+1$ 和 $c+2$ 的多项式。我们通过对 $Q(w)\ \xi(w) = P(w)$ 两边求 u 阶导数，并代入 $w = 0$，求得 $\pi(u)$ 的取值，从而得到 $\pi(u)$ 的表达式。这里， $\xi(w)$ 可解释为系数为 $a_u = \pi(u)$ 的概率生成函数，因此可以通过求 u 阶导数、令 $w = 0$并除以 $u!$ 来计算破产概率，这是概率生成函数理论中的标准步骤。

由此得到的退出概率在图 4 中针对 c 的前两个取值绘出。正如预期，在达到 $\vartheta > c/(c+1)$（即稳定性条件）之前，退出始终是确定的。先验信念参数的选取旨在确保代理人在第 0 轮不会退出。先验信念参数的影响是，对于 $c = 1$， $u_{\text{crit}} = 3$；对于 $c = 2$， $u_{\text{crit}} = 2$。

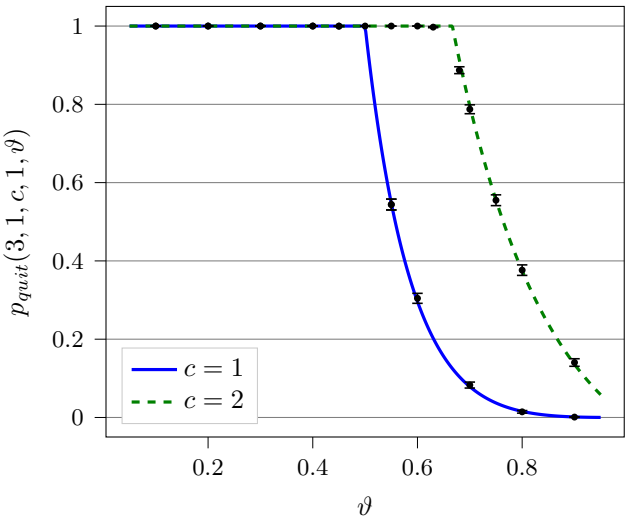


FIGURE 4. 不同 c 取值下，单个代理退出的概率随 ϑ 的变化曲线，其中 $\alpha = 3$， $\beta = 1$ 和 $r = 1$。解析结果以线条表示，模拟结果（4000 次迭代）以带有 95%置信区间的点表示。

与前文一样，我们也关注代理人给予信任并留在系统中的期望时间，条件是他们最终会退出。由于朝障碍方向的步长可能较大，这里也会遇到类似的困难；不过，在本分析中，无需增加太多额外机制即可处理这些问题。下面的引理给出了代理人给予信任直至退出的期望时间，条件是他们最终会退出。

引理 6 （ $r = 1$且 $c \in \mathbb{N}$时的预期退出时间）。 定义

$$\Phi(w,z) := \frac{(1-\vartheta)z\left(\frac{w-w^{c+1}}{1-w}\right) - \vartheta z\,\bar{\varphi}(z)}{1-(\vartheta z)/w - w^c z(1-\vartheta)}, \quad w \in (-1,1), \tag{37}$$

和 *le* t $w(z)$，对于任意给定的 $z \in (-1,1)$，记 $w \in (-1,1)$ *i* 的唯一解 n

$$w^{c+1}(1-\vartheta)z + \vartheta z = w, \tag{38}$$

and

$$\bar{\varphi}(z) := \frac{(1-\vartheta)zw(z) + \vartheta z - w(z)}{1 - w(z)}. \tag{39}$$

假设 $\vartheta > c/$（$c+1$），代理人的效用参数由 $c \in \mathbb{N}$ 和 $r = 1$ 给出。相应随机游走在 u 处到达吸收屏障的期望时间满足

$$\mathbb{E}[\tau(u) \,|\, \tau(u) < \infty] = \frac{\varphi'(1,u)}{\varphi(1,u)}, \tag{40}$$

其中

$$\varphi(z,u) := \frac{\partial^u}{\partial w^u} \frac{\Phi(w,z)}{u!}\bigg|_{w=0}. \tag{41}$$

为了将 Φ 完全表示为仅含 w、ϑ、和 z 的形式，首先利用这样一个事实：分母的根必然也是分子的根（以使 $\Phi(w,$ $z)$ 有界），并利用这一事实将 $\varphi(z,$ $1)$ 表示为含 z 的形式。附录 A 中给出的证明使用了对象 $\varphi(z,$ $u)$: $= \mathbb{E}[z^{\tau(u)}\mathbf{1}\{\tau(u) < \infty\}]$，随后将其用于生成函数

$$\Phi(w,z) := \sum_{u=1}^{\infty} w^u \varphi(z,u) \tag{42}$$

对于 $w \in (-1,$ $1)$，由此我们得到结果。

在图 5 中，我们绘出了以代理人最终退出为条件的系统期望时间的数值结果和模拟结果 q $(\alpha,$ $\beta,$ $c,$ $1,$ $\vartheta)$。我们针对实例 $c = 1,$ 2 进行计算，使用由形状参数 $\alpha = 3$ 和 $\beta = 1$ 确定的贝塔先验信念分布，即，参数选择与图 4 中使用的相同。与 $c = 1$ 和 $r \in \mathbb{N}$ 的情形一样，我们发现期望时间逐渐增加并趋近于一条渐近线，该渐近线现在位于 $\vartheta = c/($ $c+1)$。超过这个临界 ϑ 后，期望时间开始下降，因为失败越来越集中在较早的时间步。主体在系统中停留的时间越长，就越有机会从机构获得更多不带偏见的经验，因此也更有可能维持信任关系。

4. 一起信任

本节将模型从一个信任者扩展到两个信任者。为此，我们在模型中加入一种沟通机制（即信任者之间的沟通），同时保留短视理性的假设。

我们考察这样一些主体：他们彼此沟通的频率，与他们作出是否给予信任的决策的频率相同。这种交互有两种形式。第一种是主体彼此完全共享各自的交互历史。第二种是主体完全不交流各自的结果，只观察对方在每一轮采取了什么行动。

4.1. 共享建模方面。 在这两种机制中，我们都有两个主体，他们各自与机构进行交互。因此，每轮中每个主体都要从行动集 $\mathcal{A} = \{0,1\}^2$ 中作出决策，其中 $A_{i,\ t} = 1$ 若主体 $i \in \{1,2\}$ 在第 t 轮给予信任。对于每一轮 t 和每个主体 i,

$$X_{i,t} = \begin{cases} 1, \text{ with probability } \vartheta \\ 0, \text{ with probability } (1-\vartheta), \end{cases} \tag{43}$$

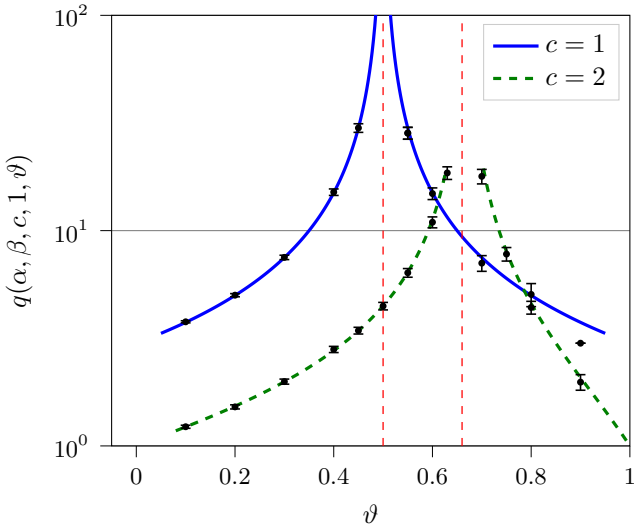


FIGURE 5. 在代理人于某个 $t < \infty$ 退出的条件下，系统中的期望时间随 ϑ 的变化图，展示了不同 c 取值下的情况，其中 $\alpha = 3$、$\beta = 1$和 $r = 1$。分析结果以曲线表示，模拟结果则以带有95%置信区间的点表示。

其中 $X_{i,\,t} = 1$ 表示该轮给予的信任随后得到了机构的兑现。与单个学习者情形类似，主体会使用

$$\widehat{S}_{i,t} = \sum_{n=0}^{t} X_{i,n} \mathbf{1}_{\{A_{i,n}=1\}}, \tag{44}$$

以记录在代理人给予信任的轮次中，信任被兑现给他们的次数。我们还将对每个代理人遭受信任背叛的次数定义为 $i = 1$，2，方法是

$$\widehat{F}_{i,t} = \sum_{n=0}^{t} (1 - X_{i,n}) \mathbf{1}_{\{A_{i,n}=1\}}. \tag{45}$$

这些主体采用与之前相同的贝塔先验信念分布。因此，代理人 i 的先验信念分布（形状参数为 α_i、$\beta_i \in \mathbb{N}$）为

$$P_{i,0} = B(\theta; \alpha_i, \beta_i) = \frac{\theta^{\alpha_i - 1}(1-\theta)^{\beta_i - 1}}{\int_0^1 y^{\alpha_i - 1}(1-y)^{\beta_i - 1} \mathrm{d}y}, \quad \theta \in [0, 1]. \tag{46}$$

为简单起见，从现在起我们只考虑齐次先验的情况（即，$\alpha_1 = \alpha_2$和 $\beta_1 = \beta_2$）。需要注意的是，非齐次先验也可以用完全相同的方式处理，但需要明确每个主体对其邻居的先验了解多少。可以假设他们准确知道邻居的先验，也可以假设他们对此持有某种信念分布；显然，这样的表述会更加繁琐，但在概念上并不复杂。

由于我们只考虑同质情形，因此省略先验信念分布的下标，简记为 $P_0(\cdot)$。将代理人 i 在第 t 轮收到的信息记为 $I_{i,\,t}(\theta)$。该信息的具体表述取决于通信方式的选择，这是建模的结果；第一种机制见 §4.2 中的（50），第二种机制见 §4.3 中的（59）。由此得到的信念更新规则为

$$P_{i,t}(\theta) = \frac{\theta^{\widehat{S}_{i,t}}(1-\theta)^{\widehat{F}_{i,t}} P_0(\theta) I_{i,t}(\theta)}{\int_0^1 y^{\widehat{S}_{i,t}}(1-y)^{\widehat{F}_{i,t}} P_0(y) I_{i,t}(y)\, \mathrm{d}y}, \quad \theta \in [0, 1]. \tag{47}$$

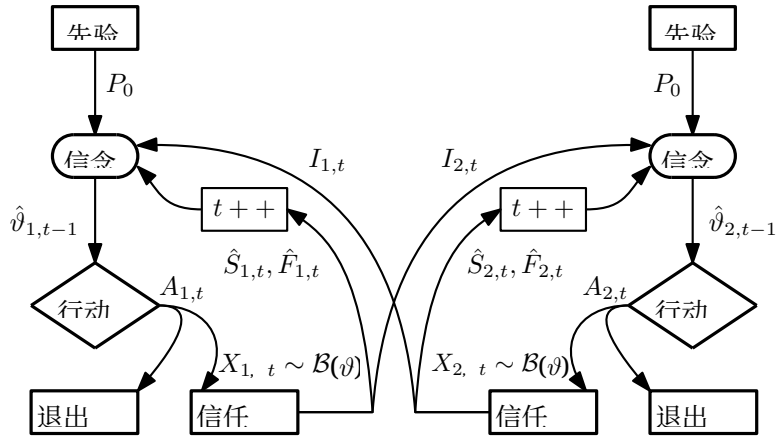


FIGURE 6. 学习的可观察奖励模型。

（47）与（2）的区别在于前者包含代理人收到的信息。随后，主体使用其信念分布的均值作为对机构真实可信度的点估计，并据此决定是否在第 $t+1$ 轮给予信任。如果我们定义

$$\widehat{\vartheta}_{i,t} = \int_0^1 \theta P_{i,t}(\theta)\,\mathrm{d}\theta, \quad \theta \in [0,1], \tag{48}$$

即，代理人 i 在第 t 轮结束时持有的平均信念，然后代理人 i 在第 $t+1$ 轮采取行动$A_{i,\ t+1}$ 其中

$$A_{i,t} = \begin{cases} 1, \text{ if } r\widehat{\vartheta}_{i,n} - c(1-\widehat{\vartheta}_{i,n}) \geq 0, \quad \forall n \in \{0,1,\ldots,t-1\}, \\ 0, \text{ otherwise.} \end{cases} \tag{49}$$

请注意，条件（49）可以改写为 $\vartheta_{i,\ n} \geq \theta_{\text{crit}}, \quad \forall n \in \{0,1,\ldots,t-1\}$，其中$\theta_{\text{crit}} := c/(c+r)$。接下来的两个小节将更详细地讨论这两种不同的沟通机制。

4.2. 可观测奖励（OR）。这种情形考虑的是主体之间就各自与机构打交道的经历进行充分沟通的情况。该模型的动态在图6中以图示形式呈现。与单学习者模型（为便于比较，在图1a中表示）相比，该模型的扩展在于主体之间共享信息，其中包含了机构对所给予信任的回应。

因此，每个代理人都拥有相同的可用信息，并且在同质的先验信念下，他们对真实可信度的估计值相同。在这种情况下，我们定义代理人 i 在轮次 t中从其邻居 $j \neq i$ 接收的信息，

$$\theta \mapsto I_{i,t} := \theta^{\widehat{S}_{j,t}}(1-\theta)^{\widehat{F}_{j,t}}, \quad \theta \in [0,1]. \tag{50}$$

我们看到，该信息对每个代理人而言恰好包含另一代理人的交互历史，因此在图6中，我们有一条从代理人 i的“信任”行动指向代理人 j的“信念”分布的线。

如前所述，主体对真实可信度的估计值相同。因此我们可以从 $\vartheta_{i,t}$ 中去掉下标 i，并注意到它由下式给出

$$\widehat{\vartheta}_t = \frac{\alpha + \sum_{i=1}^2 \widehat{S}_{i,t}}{\alpha + \beta + \sum_{i=1}^2 \widehat{S}_{i,t} + \sum_{i=1}^2 \widehat{F}_{i,t}}. \tag{51}$$

与单代理模型类似，定义 $Z_t = c\sum_{i=1}^{2}\widehat{F}_{i,t} - r\sum_{i=1}^{2}\widehat{S}_{i,t}$后，我们得到一个随机游走，其步长为

$$Z_{t+1} = \begin{cases} Z_t + 2c, & \text{with probability } (1-\vartheta)^2 \\ Z_t + c - r, & \text{with probability } 2\vartheta(1-\vartheta) \\ Z_t - 2r, & \text{with probability } \vartheta^2, \end{cases} \tag{52}$$

在 $u = r\alpha - c\beta + 1$ 处设有吸收屏障。

在先验具有异质性的情况下，该模型的表述不会发生显著变化。主体的命运不再像之前那样相互关联，因此将简单地定义两个如（52）所示的随机游走，每个主体各一个。每个随机游走都有其自身的吸收屏障，由 $u_i = r\alpha_i - c\beta_i + 1$ 定义。回想一下，α_i，β_i，r 以及对于所有 i 而言的 c 均取自 $\mathbb{N}$，因此 u_i 对两个 i 而言也都是整数。如果在某个时刻 t_q 其中一个主体退出，那么剩余主体将按照单个代理模型的动态继续运行，并采用以下定义的新转移：

$$Z_{t+1} = \begin{cases} Z_t + c, & \text{with probability } (1-\vartheta) \\ Z_t - r, & \text{with probability } \vartheta, \end{cases} \tag{53}$$

对于 $t > t_q$。

4.3. 可观察行动（OA）。在这种情境下，主体不会通过分享有关交互历史的信息来进行明确沟通。相反，他们会观察邻居的行动，并据此推断邻居可能经历过哪些历史，从而导致这样的结果。微妙复杂之处也由此产生：理性的主体需要牢记，邻居会观察自己的行动，因此自己的行动也会影响邻居所做的决策。

在此模型中，两个主体之间的通信程度是二元的，表示在第 t 轮次结束时是否给予了信任。这些信息仅纳入主体第 $t+1$ 轮次的信念分布中。图 7 以图示方式呈现了此模型。该模型的关键在于一条从一个主体的“行动”决策指向另一主体“信念”分布的连线。不给予信任这一行动意味着该主体将永久退出系统。退出的主体不会继续根据邻居的行动更新其信念。退出主体发送的信息，在未退出主体留在系统期间，会被用作未退出主体接收的信息。

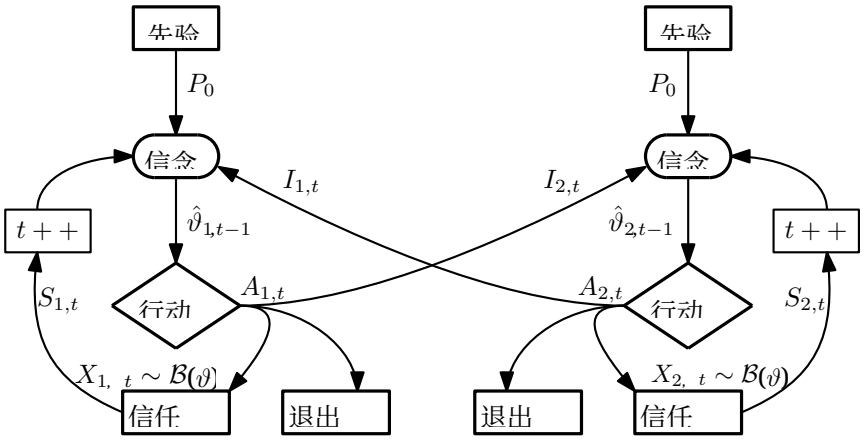


FIGURE 7. 可观察行动学习模型。

通信的性质要求分别定义主体给予信任时发送“是”信号所携带的信息，以及主体不给予信任时发送“否”信号所携带的信息。因此，我们对 $i = 1,2$, $j \neq i$ 和 $t \in \mathbb{N}$ 定义如下，

$$I_{i,t} := \begin{cases} Y_{i,t} & \text{if } t < \tau_j, \\ N_{i,\tau_j} & \text{if } t \geq \tau_j, \end{cases} \tag{54}$$

其中 τ_j 是代理人 j 为 $j = 1$退出的轮次，2。我们分别在 (59) 和 (61) 中正式定义 $Y_{i,\,t}$ 和 $N_{i,\,t}$。根据理性假设，主体会使用二元信息，并且知道自己的出站信息直到第$t-1$轮都被另一个主体理性地使用。借助这些信息，他们可以推断出一系列可能的历史序列，这些历史序列可能导致邻居做出相应决策。

我们定义信念分布 R_t^i，其计算仅考虑代理人 i收到的报告（由代理人 j 在时间 t 发送给代理人 i的信号）：

$$R_t^i = \frac{P_0(\theta) I_{i,t}(\theta)}{\int_0^1 P_0(y) I_{i,t}(y)\,\mathrm{d}y}, \quad \theta \in [0,1]. \tag{55}$$

接着，我们定义由 R_t^i 和$\theta^x(1-\theta)^{t-x}$ 组合得到的辅助信念分布，其中 x 是产生 $\mathcal{D}_t(x)$ 这一分布的可能正向经历次数之一：

$$\mathcal{D}_t(x) = \frac{\theta^x (1-\theta)^{t-x} R_t^i}{\int_0^1 y^x (1-y)^{t-x} R_t^i\,\mathrm{d}y}, \quad \theta \in [0,1]. \tag{56}$$

这使我们能够定义代理人 i的允许位置集合， $\mathcal{J}_t^i$ 为

$$\mathcal{J}_t^i = \left\{ x \in \{0,1,\ldots,t\} \,\middle|\, \mathbb{E}_{\theta \sim \mathcal{D}_t(x)}[\theta] \geq \theta_{\text{crit}} \right\}. \tag{57}$$

对于所有 $x \in \mathcal{J}_t^i$， $\mathcal{D}_t(x)$ 都是代理人 i持有的可能信念分布。随后，在 $I_{j,\,t}$中对这些可能的信念分布求和，并根据 $w_x(t-1)$ 加权，即依据这些 x 次正向经历可能实现的方式数目加权（即，截至时间 t，产生 x 次正向经历的历史序列数）。递归定义 $w_x(t)$ 如下：

$$w_x(t) = \begin{cases} w_x(t-1) + w_{x-1}(t-1) & \text{if } (x-1) \in \mathcal{J}_{t-1} \\ w_x(t-1) & \text{if } (x-1) \notin \mathcal{J}_{t-1} \\ 1 & \text{if } x = t \\ 0 & \text{if } x > t \text{ or } x \notin \mathcal{J}_{t-1}, \end{cases} \quad \text{for } x \in [0,t] \text{ and } t \in \mathbb{N} \tag{58}$$

在初始条件为 $w_x(1) = 1$ 对所有 $x \in \mathcal{J}_1$均成立时。代理人 i 从代理人 j 处收到的信息，考虑到代理人 j在第 t 轮的决策是给予信任，定义如下

$$\theta \mapsto Y_{i,t} := \sum_{x \in \mathcal{J}_{t-1}^j} w_x(t-1)\, \theta^x (1-\theta)^{t-1-x}, \quad \theta \in [0,1]. \tag{59}$$

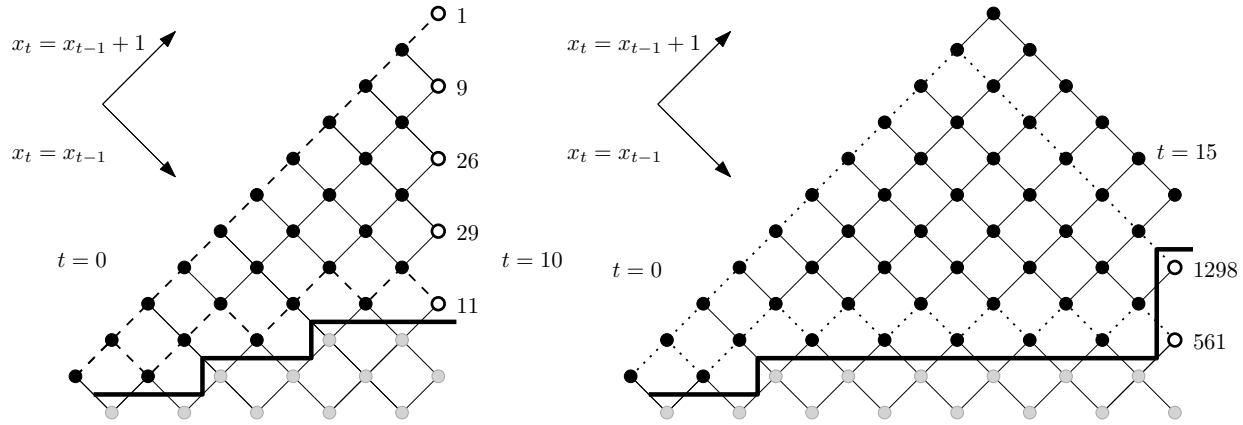
我们可以将 $w_x(t)$ 理解为从 $(0,\ 0)$ 到 $(x,\ t-x)$ 的路径数量，这些路径在 $\tau = 1$、2、.....、t 时均保持 $x_\tau \in \mathcal{J}_\tau$。图 8a 展示了“是”信号对应的允许的路径，其中进出 $\mathcal{J}_t$ 的边界由粗水平线表示。

为了定义随“否”信号发送的信息，我们定义这样一个数值范围：在第 $t-1$ 轮之前会产生“是”信号，并在第 t 轮产生“否”信号：

$$\mathcal{K}_t^i = [\max\{0, \inf_x \mathcal{J}_{t-1}^i\}, \ldots, \inf_x \mathcal{J}_t^i - 1]. \tag{60}$$

接下来，我们来看通过“否”信号传递的信息：

$$\theta \mapsto N_{i,t} := \sum_{x \in \mathcal{K}_t} w_x(t)\, \theta^x (1-\theta)^{t-1-x}, \quad \theta \in [0,1] \tag{61}$$



(A) $w_x(t)$计为“是”信号。

(B) $w_x(t)$计为“否”信号。

FIGURE 8. 对观察到的行动进行解释时所使用的加权 $w_x(t)$示意图。

其中， $w_x(t)$ 的定义如上。在时间 t 发出“否”信号时，允许的随机游走数量，就是在时间 $t-1$ 发出“是”信号的允许随机游走数量中，距离退出还差一次负面经历的那些路径的数量。图 8b对此进行了说明：水平线表示 x 的内外边界，点线表示 x 的内外边界。请注意，后续轮次中（即首次发出“否”信号之后）“否”信号传递的信息不会改变，正如我们在（54）中所提到的。

主体最终将自身信息与信号信息相结合，得出一个估计值：

$$P_{i,t}(\theta) = \frac{\theta^{\widehat{S}_{i,t}}(1-\theta)^{\widehat{F}_{i,t}} P_0(\theta) I_{j,t}(\theta)}{\int_0^1 y^{\widehat{S}_{i,t}}(1-y)^{\widehat{F}_{i,t}} P_0(y) I_{j,t}(y)\,\mathrm{d}y}, \quad \theta \in [0,1], \tag{62}$$

并在第 $t+1$ 轮中于 $\mathbb{E}_{\theta\sim P_{i,\,t}(\theta)}[\theta] \geq \theta_{\text{crit}}$ 时给予信任。

4.4. 所呈现模型之间差异的说明。在本小节中，我们为两位参与者展示一组示例经历。这些示例用于说明 OA模型和 OR模型中的动态如何运作。

我们考察参数为 $c = 2, r = 1$ （因此 $\theta_{\text{crit}} = 2/3$）和 $u_{\text{crit}} = 2$ （$\alpha = 5, \beta = 2$）的双主体模型。暂且不判断这些交互结果是否实际出现，假设机构对这两个主体在前四轮（X_i,1, X_i,2, X_i,3, X_i,4） =(0,0, 1, 1) 中的回应适用于 $i = 1$，而在（X_i,1, X_i,2, X_i,3, X_i,4） = (1, 1, 1, 0) 中的回应适用于 $i = 2$。图 9 展示了单智能体模型以及可观察奖励模型中的相应随机游走解释。

我们观察到，在单个代理模型中，第一个代理人在第一次交互后就会退出，但在 OR 模型中，他们会继续行动，不会在此处停止。OA 模型不太容易通过视觉方式呈现，因此需要进一步阐释。我们在表 3 中汇总了代理人1和代理人2各自的单个代理模型以及双主体模型中各变量随时间的变化。请注意，在 OR 模型中，代理人1和代理人2的估计值始终相同，因为他们从相同的先验开始，并且每轮都能获取相同的信息。第一个代理人的估计值在第一轮后不再变化，这是因为他们退出了；在单个代理模型和 OA 双主体模型中，情况也是如此。这是因为，当他们决定不予信任时，

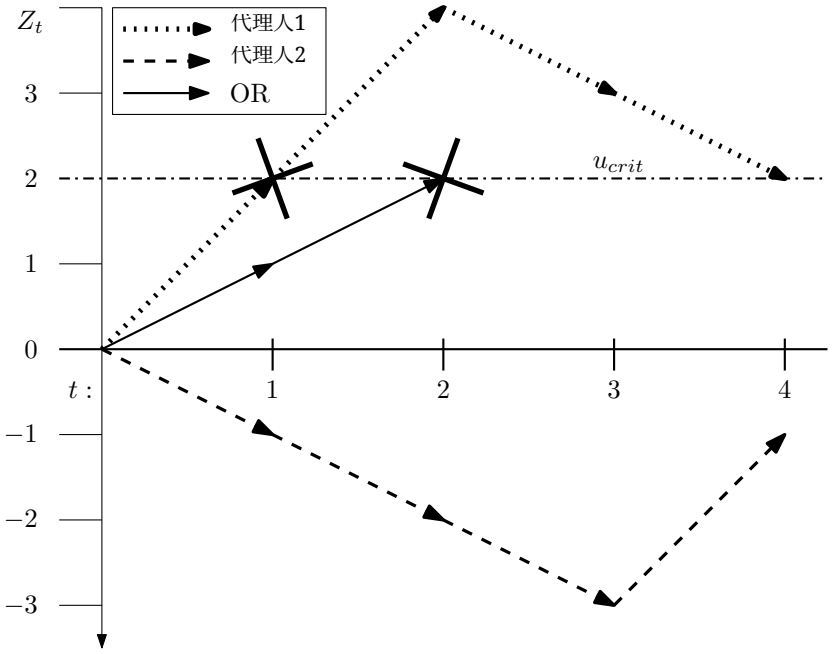


FIGURE 9. OR模型的随机游走解释，以及相应主体的个人样本路径。

他们还没有收到包含代理人2信息的信号。因此，他们并未观察到机构对 $X_{i,\ t}$ 针对 $i = 1$ 和 $t \geq 2$ 的反应，也没有将这些反应用于更新代理人1的信念。

TABLE 3. 估计值，ϑ_t 在相应模型中。

t	代理人1	代理人2	或（1）	或（2）	OA（1）	OA（2）
0	0.714	0.714	0.714	0.714	0.714	0.714
1	0.625	0.750	0.667	0.667	0.625	0.750
2		0.778	0.636	0.636		0.700
3		0.800				0.727
4		0.727				0.667

在 OA模型的第一轮中，两个主体的行为与单智能体模型中的行为完全相同，因为它们的首次行动不包含信息。第一轮结束后，第一个代理人退出，并在第 2 轮及所有后续轮次中采取“不信任”行动。这向第二个代理人表明，代理人1在第一轮滥用了信任。在这种情况下$w_x(1) = 1$（作为 w_x 的初始条件的一部分）以及 $I_{2,\ 2} = \theta^0\ (1-\theta)^{\ 1}$ 而其个人S_2 等于 2，这提供了

$$P_{2,2}(\theta) = \frac{\theta^2 \cdot \theta^{5-1}(1-\theta)^{2-1} \cdot (1-\theta)^1}{\int_0^1 y^2 \cdot y^{5-1}(1-y)^{2-1} \cdot (1-y)^1 \mathrm{d}y}. \tag{63}$$

该信念分布的均值被用作点估计；它是

$$\widehat{\vartheta}_2 = \int_0^1 \theta \frac{\theta^6(1-\theta)^2}{\int_0^1 y^6(1-y)^2\,\mathrm{d}y} \mathrm{d}\theta = 0.7. \tag{64}$$

为便于说明，我们还展示代理人2在第3轮采取行动时发送的信息，也就是，$I_{1,\ 3}$。首先，他们构建 R_2^2，即综合考虑以下内容后形成的信念

代理人2截至第2轮收到的报告：

$$R_2^2(\theta) = \frac{P_0(\theta) I_{2,2}(\theta)}{\int_0^1 P_0(y) I_{2,2}(y)\,\mathrm{d}y} = \frac{\theta^4(1-\theta)^2}{\int_0^1 y^4(1-y)^2\,\mathrm{d}y}. \tag{65}$$

随后构造辅助分布 $\mathcal{D}_2$：

$$\mathcal{D}_2(x) = \frac{\theta^x(1-\theta)^{2-x} R_2^2(\theta)}{\int_0^1 y^x(1-y)^{2-x} R_2^2(y)\,\mathrm{d}y} = \frac{\theta^{x+4}(1-\theta)^{4-x}}{\int_0^1 y^{x+4}(1-y)^{4-x}\,\mathrm{d}y}. \tag{66}$$

然后利用该分布构造集合 $\mathcal{J}_{t=2}^{i=2} = \{2\}$，因为只有 $\int_0^1 \mathcal{D}_2(x)\,\mathrm{d}\theta \geq 2/3$ 才满足 x。考虑到 $x = t$，我们知道 $w_2(2) = 1$，这意味着 $I_{1,\ 3} = \theta^2$。

这个例子展示了信息级联：除 OR模型外，在任何模型中，第二个代理人在前 4 轮都不会退出。另一方面，对第一个代理人的影响是，在单智能体模型和 OA模型中，他们会在第 1 轮退出，而在 OR模型中则会在第 2 轮退出。

5. Setup of simulation experiments

本节介绍模拟实验的设置。我们在第 6 节呈现这些模拟实验的结果，并在第 7 节讨论和解释实验结果。实验的主要目标是，从退出概率和预期退出时间（以退出为条件）两方面，评估双代理人情形下的两种通信机制。

5.1. 选择 c 和 r。 为了了解模型参数如何影响主体退出的概率（及时间），我们选择了以下 5 组（c、r）组合：(1、1)、(3、2)、(2、3)、(2、1)、(1、2)，其中 $c = r$ 是“基本情形”，而 $c = 2r$ 和 $r = 2c$是“极端端点”。选择这一比值范围有两个原因。首先，如果比值在任一方向上进一步扩大，模拟结果就会更难获得。随着 c增加，临界可信度 θ_{crit} 会向 1 移动，这意味着只有一小段 ϑ 的退出概率小于 1。在这一小范围内，退出概率会急剧下降，因为当 ϑ 非常大时，退出概率很低，这是因为退出需要滥用信任。另一方面，随着 r增加， θ_{crit} 会降低，因此越来越多的参数设定必须在已经运行缓慢的模拟中经历相对较长的时间区间。这是因为在用于构建 $\mathcal{J}_t^i$、 $\mathcal{D}_t$ 和 R_t^i 的可观察行动模型中，对于每个 $i = 1$、2都需要进行大量数值积分。 第二个原因是，这涵盖了足够广泛的成本收益比，尤其是在信任问题[3]的背景下。

5.2. 选择 α 和 β。 </style}直观而言，参数 α 和 β 可以视为我们所考察的动态发生之前与该机构进行交互的次数，在这些交互中，信任得到兑现 α 次，遭到辜负 β 次。这意味着，与 β相比， α 越大，主体的先验信念分布就越偏向 ϑ的较大值。同样，与 β相比，若 α 较小，该先验信念分布就越偏向 ϑ的较小值。先验信念参数设置 α 和 β，与

[3]将交互描述为向机构支付效用 x 时，也会产生相同的动态。此后，机构的诚信行为就相当于获得效用 $x + \delta$。另一方面，滥用行为则意味着 x 会损失。现在， $c = x$ 和 $r = \delta$，我们的范围涵盖从 $\delta = 2x$ 到 $\delta = x/2$的数值，我们将其视为信任问题。我们认为 $\delta > 2x$ 更接近赌博问题，而 $\delta < x/2$ 更接近将效用从一种形式转换为另一种形式（也就是说，不涉及任何风险）。

对 c和 r的选择决定了该实例的 $u_{\text{crit}} = r\alpha - c\beta + 1$。借用单智能体模型中单个代理的一维随机游走这一类比， u_{crit} 的值代表距离退出的“距离”。若要使对 u_{crit}的解释保持一致，我们需要进行换算：取 $u^* = u_{\text{crit}}/c$，即单智能体模型中代理人最早可以退出的时间。模型在 $c > r$和 $c < r$两种情形之间存在不对称性。在 $c > r$的情形下，代理人可能在第一轮经历信任得到兑现，但如果辜负信任的步长满足 $c > u_{\text{crit}} + r$，仍会在第二轮后退出。在 $r > c$ 的情形下，这种情况不可能发生。我们选择 α 和 β的取值，使其与 r 和 c的取值结合后，覆盖范围 $u^* = 1$，2，3。选择在 $u^* = 3$ 处停止，是为了使退出概率足够高，从而便于高效模拟。

5.3. **迭代次数与模拟时长**。 在我们的仿真研究中，我们的目标是对所研究的参数得出可靠的估计值。我们总共针对 4 000 个信任主体运行每个模型，即，单智能体模型运行 4 000 次，而双智能体模型运行 2 000 次。每次模拟的最长时长为 500 个时间步；对于包含 $\vartheta \in \{0.84, 0.9\}$ 的模拟，我们则将模型的最长运行时长设为 200 个时间步。这些选择既考虑了置信区间的宽度，也考虑了执行的可行性。鉴于双智能体模型中的动态会加快（因为每个时间步可获得更多信息），这段时间对于两个双智能体模型而言也应当足够。

在具有最高 ϑ 的交互中，退出概率很低，且集中在最初几个时间步。我们可以审慎地（偏悲观地）估计，为避免遗漏任何相关的“概率质量”，需要多少个时间步。为此，我们将马尔可夫不等式应用于单个代理模型中的预期退出时间。举例来说，考虑参数设定为 $c = 2$， $r = 1$， $\alpha = 5$，且 $\beta = 2$ （使得 $u^* = 1$）的情形，其参数设定为 $\vartheta = 0.84$。在单个代理模型中，预期退出时间为 $\mathbb{E}[\tau \,|\, \tau < \infty] \approx 3.17$。根据马尔可夫不等式，我们知道在时间 $\tau \geq 200$ 退出的概率有如下界限：

$$\mathbb{P}(\tau \geq 200 \,|\, \tau < \infty) \leq \frac{\mathbb{E}[\tau \,|\, \tau < \infty]}{200} = 0.01587. \tag{67}$$

因此，在所进行的 N 实验中，我们预期错过退出事件的次数为 $L = N \cdot \mathbb{P}(\tau \geq 200 \,|\, \tau < \infty)$ $\cdot\, \mathbb{P}(\tau < \infty)$。取 $N = 4\,000$，得到以下界限

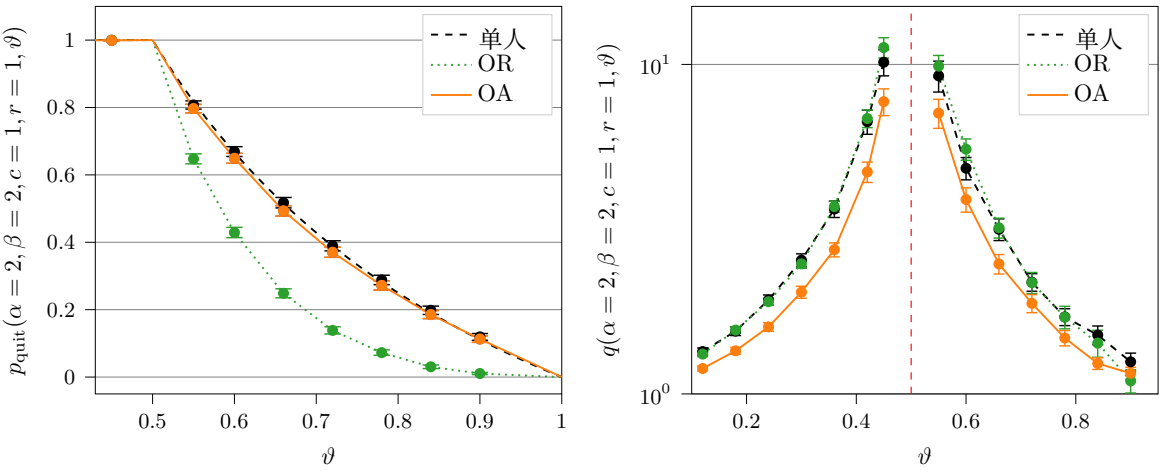
$$L < 4000 \times 0.01587 \times 0.2635 = 16.727. \tag{68}$$

单学习者情形是一个保守的基准，因为与双代理模型相比，其通信较少，因此我们预计双智能体模型中的主体会更早退出。这一估计值的保守程度可以从以下事实看出：在单智能体模型中（针对所述参数设定），最晚退出发生在第 45 轮；在双智能体 OA 模型中，最晚退出发生在第 40 轮；而在双智能体 OR 模型中，最晚退出发生在第 17 轮。

6. Results of experiments

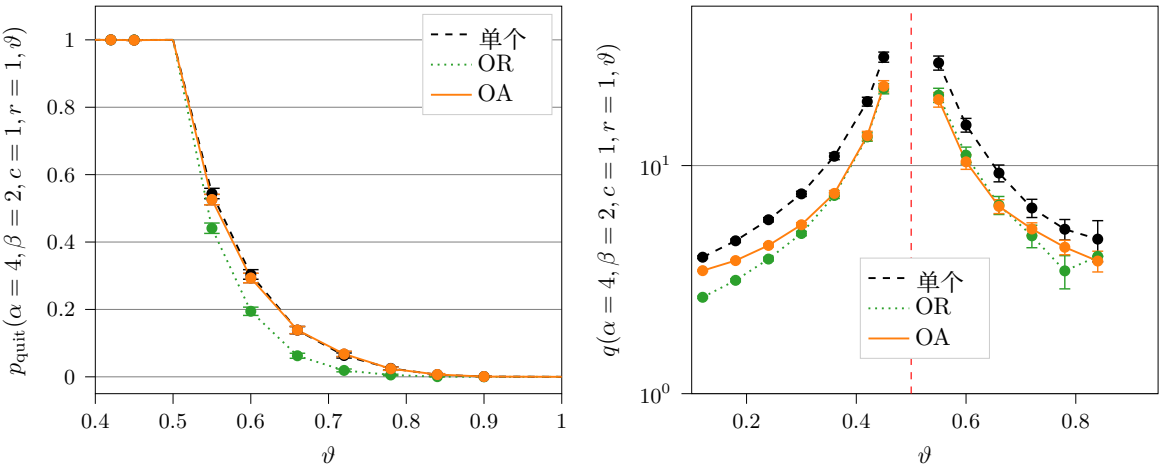
本节展示 §5 中所述实验的结果。我们绘制了 (c, r) 和 u^* 取特定值时的结果，见图 10a–15b；所有实验的数值输出以表格形式列于附录 D。图中涵盖的设置为 (c, r) =(1, 1),(2,1),(1,2) 和 $u^* = 1, 3$；其中，我们将合成点 $p_{\text{quit}}(\alpha, \beta, c, r, \vartheta = \theta_{\text{crit}})$ =1 和 $p_{\text{quit}}(\alpha, \beta, c, r, \vartheta = 1) = 0$ 添加到图中的曲线上，以展示超出模拟能力范围的动态变化。为了能够观察到细微差异，

我们主要展示退出未必会发生的区域内的退出概率。



(A) 退出概率 (B) 预期退出时间

FIGURE 10. $u^*=1$ 情况下 $c=r=1$ 的模拟结果。



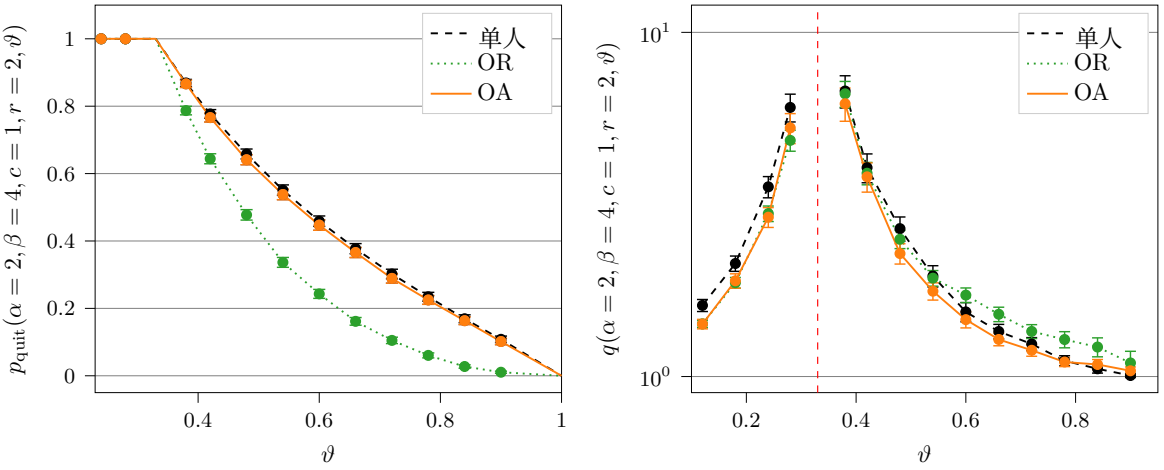
(A) 退出概率 (B) 预期退出时间

FIGURE 11. 在 $u^*=3$ 情形下的 $c=r=1$ 模拟结果。

6.1. 退出概率。在参数区间 $\vartheta < \theta_{\text{crit}}$中，退出概率应保持为1：与单智能体模型相比，双主体模型中的主体可获得相同甚至更多的信息。表5a、6a和7a中的0.999这一数值本可通过增加模拟工作量来改善（即，增加运行次数并采用更多时间步）。

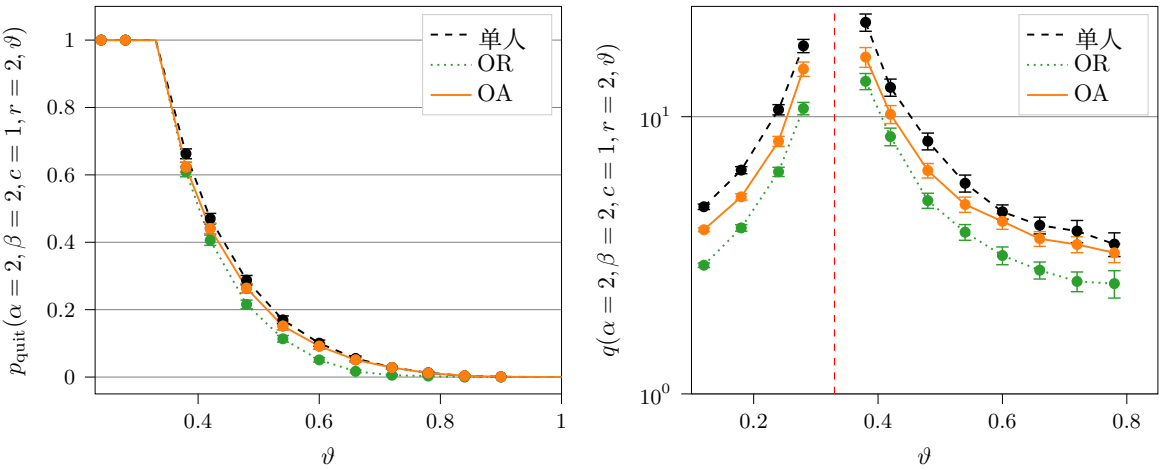
在（更有趣的）参数区间 $\vartheta > \theta_{\text{crit}}$中，对于大多数参数设定，退出概率从OR模型到OA模型，再到单智能体模型逐步增加。与 $u^*=2$ 和 $u^*=3$相比，这一趋势在 $u^*=1$ 中更为明显。表6a和7a中存在两个与此趋势相反的例外：对于 $\vartheta=0.65$，在参数设定$c=3$ 和 $r=2$下为0.66，且 $u^*=2$，3。不过，我们观察到原始置信区间存在重叠。

为了确定这种差异是否具有显著性，我们针对较快的情形（ $u^*=2$）进行了更多模拟。对于计算开销较小的 OR 模型，我们进行了



(A) 退出概率 (B) 预期退出时间

FIGURE 12. $c = 1$ 和 $r = 2$ 在 $u^* = 1$ 情况下的模拟结果。

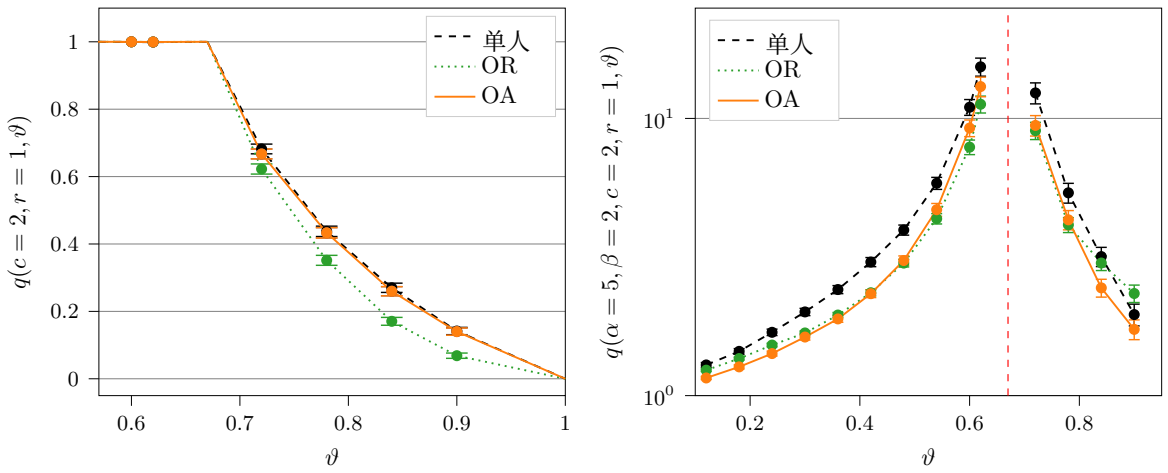


(A) 退出概率 (B) 预期退出时间

FIGURE 13. $u^* = 3$ 情形下 $c = 1$ 和 $r = 2$ 的模拟结果。

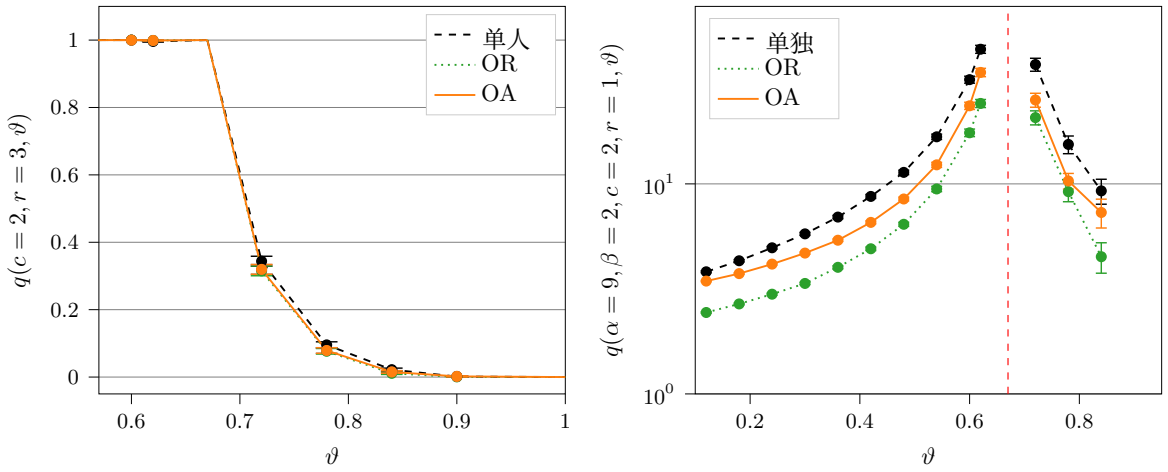
总计 200 000 次运行。对于速度较慢的 OA模型，我们总共进行了 40 000 次模拟。图16 展示了点 $\vartheta = 0.65$，0。66 的置信区间。我们观察到一个或许出乎意料的结果：在这一参数设定下，OA模型的退出概率更低。因此，在 $\vartheta > \theta_{\text{crit}}$ 所处的参数区间内，就退出概率而言，OR模型并非总是优于 OA模型。

OA模型中退出概率较低的原因之一，与通信发生的时间有关，即在轮次结束时 。在OR 模型中，代理人一旦获得自己的奖励，也会观察邻居的奖励。换言之，每轮会获得两条信息片段。然而，在OA模型中，代理人只有在已经完成该轮自己的行动后，才会同时观察邻居的行动。这意味着，代理人首次在第 t 轮观察到邻居没有给予信任时，只能利用其中包含的信息来指导自己在第 $t + 1$轮的行动。与此同时，由于他们在第 t轮尚未退出，因此在必须再次做出决策之前，至少还能得知该轮的结果。我们在附录 C 中进一步阐述这一影响。



(A) 退出概率 (B) 预期退出时间

FIGURE 14. $c=2$ 和 $r=1$ 在 $u^*=1$ 情况下的模拟结果。



(A) 退出概率 (B) 预期退出时间

FIGURE 15. $u^*=3$ 情况下 $c=2$ 和 $r=1$ 的模拟结果。

6.2. 预期退出时间。 总体而言（在参数区间 $\vartheta < \theta_{\text{crit}}$内），双主体模型的预期退出时间低于单智能体模型。此外，OR模型的预期退出时间往往最低（即效果最大），不过与OA模型相比，两者的结果有相当多的交叉。这一趋势的主要例外出现在 $c=r=1$且 $u^*=1$时，如图10b所示。在这种设置下，OA模型的表现优于另外两个模型，而后两者的退出时间相近。此外，在图14b所示的$c=2$、 $r=1$ 和 $u^*=1$ 情况下，以及表5b所示的 $c=3$、$r=2$ 和 $u^*=1$ 情况下，也存在例外。

在 $\vartheta < \theta_{\text{crit}}$ 参数区间内，存在这样一种趋势（但也有例外）：OR模型中的退出早于另外两个模型。在OR模型中，主体在每个互动回合中都会接触到两个结果，因此所获得的无偏信息是单智能体模型的两倍。在这一参数区间内，如此获得的无偏信息很可能表明该机构值得信任，因此随着时间推移，退出变得不太可能（即，退出必须迅速发生，否则就不会发生）。当$u^*=3$ 时，这一趋势最为明显，如图13b和图15b以及表7b所示。

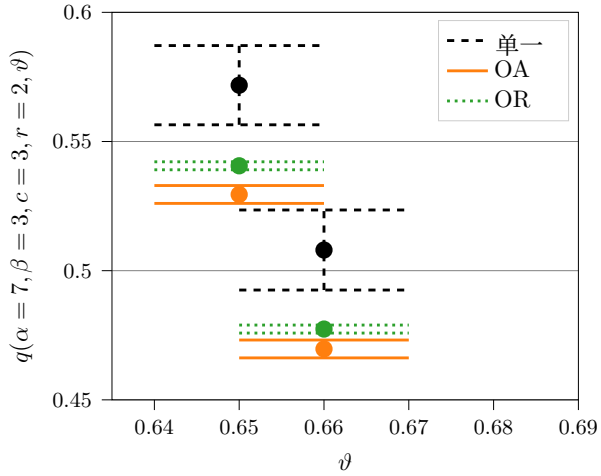


FIGURE 16. 关于退出概率的额外模拟运行结果。在这些模拟运行中，$c = 3$、$r = 2$和 $u^* = 2$。

7. DISCUSSION

本节将从数值实验中得出结论，讨论这些发现与文献中相关研究的联系，并提出未来研究的方向。

7.1. 一般性结论。 在所有模型中，预期退出时间都呈现出相似的模式：随着 ϑ 增加，预期退出时间会延长，直到 θ_{crit}，之后则会缩短。具体而言，这意味着客户在某机构的平均任职时间较长，并不代表该机构值得信任；换句话说，仅凭这一信息，无法判断它们处于临界值的哪一侧。某机构的可信度可能刚好足以让客户继续信任它，从而维持这段关系，直到客户从中获得正的净效用，但实际上又不足以保证这段关系无限期持续。另一方面，若某机构仍维持着许多关系，同时只有少数关系在很短时间内结束，这或许能很好地表明该机构确实值得信任。

通过比较相同 c 和 r 在不同 u^* 下的退出概率（例如图 10a 和 11a），我们看到，对于 $u^* = 2$，3，三个模型之间的差异相对较小。这表明，乐观地开始能够让主体长期获得信任值得信赖的机构的良好机会。

我们发现了一个（或许）令人瞩目的现象：两种不同的双代理人沟通机制所产生的效果并非单调。这一点在以下情形中尤为明显：（a） $c = r = 1$ 和 $u^* = 1$的情形，如图10a和10b所示；（b） $c = 3$、$r = 2$和 $u^* = 2$的情形，如图16所示。在（a）中，就信任值得信赖的机构的概率而言，OR模型优于OA模型。然而与此同时，在决定何时结束与不值得信赖的机构的信任关系方面，OA模型比OR模型更早做出决策。在（b）中，当 $\vartheta = 0.65$，0。66时，就信任值得信赖的机构的概率而言，OA模型优于OR模型。这些结果表明，我们不能断言OR模型总是“优于”OA模型。事实上，这取决于参数设定，也取决于你认为哪项标准最重要：从长远来看做出正确的决策，还是确保相对迅速地结束与不值得信赖的机构的关系。

这一出人意料的结果部分源于底层动态的“时序”。主体同时做出是否信任的决策，因此只能在下一轮使用来自邻居行动的信息。我们在附录 C 中通过一个两轮次的情形对此进行了说明。

我们总结数值结果所带来的启示：

- 通信总是有帮助的：与单个代理模型相比，通信提高了永远不再信任一家值得信赖的机构的概率，并缩短了与一家不值得信赖的机构保持关系直至退出的预期时间。
- OA模型可能优于或劣于OR模型，这取决于所关注的绩效指标（退出概率或预期退出时间）。在某些情况下，较少的信息反而有利于主体。
- 如果没有社交网络，想要从长远来看提高信任一家值得信赖的机构的机会，一个好的方法是从更乐观的先验开始。

7.2. 反思与背景。 在本段中，我们从概念层面将我们的模型与文献中的模型进行比较。在接下来的段落中，我们将把模型的结果与文献中的结果进行比较。我们提出了一个信任模型，其中既包括通过与机构互动进行学习，也包括通过主体之间的通信进行学习。从主题上看，这项工作与社交网络信任相关，其重点在于信任可能遭受的损失。从方法论上看，这项工作与社交网络学习相关。我们的模型中的主体正在了解环境（受托人的可信度）。我们的模型以一种自然的方式拓展了这两个研究方向中的思想：主体所采取的行动与其随后用于学习的信号之间的依赖结构，以一种现实的方式拓展了社交网络学习文献中的研究；该文献通常未涵盖这一复杂性。最后，通信的可观察行动模型使主体之间的关系更加贴近现实，而社交网络信任文献中更常见的通信可观察奖励模型并不具备这一点。

我们的研究结果与 Buskens [26], 的研究结果相似，表明在信任主体分享有关受托人的信息的模型中，信任会增加。然而，在我们的模型中，这一效应还延伸至这样的模型：受托人在某一轮次中违背信任，并不立即表明其不值得信任。此外，我们将信任主体之间的通信类型扩展到了完全信息共享之外。我们发现，尽管主体通信的积极效应在完全信息共享中更强，但在通信仅限于观察行动的模型中，这种效应同样存在。因此，即使个人并未与同伴就该机构进行广泛讨论，仅仅观察其行动也能带来显著益处；此前社交网络信任文献已表明，更广泛的通信也具有这种益处 [26, 27, 28]。

在我们考察的两种通信模型中，对于所有参数设定，我们都观察到，当理性地选择信任时，理性使用社交网络信息会提高主体学习真实可信度的概率。我们还观察到，在两种通信模型和所有参数设定下，预期退出时间都会提前；当理性地选择退出时，这一点尤其有益。我们的模型对主体间的通信以及信号对行动的依赖作出了现实的假设，并表明主体间的通信有助于学习和信任。独立信号与行动这一过于简单的设定无法直接应用于信任与学习模型。信任问题的本质在于，必须将资源交由受托主体支配，才能观察其如何回应。然而，与社交网络学习文献（参见 [14, 18]）一样，我们的研究表明，更多的通信会带来更快的动态，

但是，由于信号取决于行动，有时最终会收敛到次优行动（即不选择信任）。

7.3. 未来工作。 从所构建的模型中可以看出，主体彼此通信是有益的。然而，在某些情况下，主体原本可能会继续给予信任，却因为听到邻居的说法而“错误地”停止给予信任。整体动态似乎并未受到这种影响的主导。于是，一个自然的问题是：当模型中的主体超过两个时，这种有益影响是否依然存在？目前的模型仅限于两个主体，部分原因是所需的计算工作量巨大：OA模型中的主体每轮都要进行大量数值积分，以判断哪些经历历史序列与收到的邻居通信相符。如何以可扩展的方式完成这些计算，仍是一个悬而未决的问题。另一种方法是，在主体解读邻居行动时放宽对其理性的要求，从而绕过这一问题。这样便可研究更多主体，考察它们如何从私人信号中学习，同时也要付出简化模型的成本。

未来工作的另一个方向是研究两个信任者主体之间的非对称沟通。其中一个信任者主体可以建模为新闻媒体，只发送信息而不接收任何信息。同样，一个主体也可能通过散布有关该机构的错误信息而成为恶意主体。这种主体间的不对称性或许能让技术性结果更容易实现。例如，可以以唯一发送信息的信任者的私人信号为条件，并观察接收信任者的动态。纯发送型信任主体的动态恰好遵循我们在单主体信任模型中所提出的动态。也可以通过事先决定恶意行为者将发送什么信号来对其进行建模；或者为其设定一个真实性参数，使信任主体以概率 η 传达其真实经验，并以概率 $1-\eta$传达信任遭到滥用这一信息（无论事实是否如此）。在这种模型中，诚实主体可能不仅需要了解该机构的可信度，还需要了解其从网络中收到的信息。

References

[1] N. Bajos, A. Spire, L. Silberzan, A. Sireyjol, F. Jusot, L. Meyer, J. E. Franck, and J. Warszawski, “当政府和科学家缺乏信任加剧了COVID-19疫苗接种中的社会不平等”，公共卫生前沿，2022年。[2] J. V. Lazarus, K. Wyka, T. M. White, C. A. Picchio, K. Rabin, S. C. Ratzan, J. P. Leigh, J. Hu, and A. El-Mogandes, “利用2021年23个国家的数据重新审视全球COVID-19疫苗犹豫”，《自然通讯》，第13卷，第3801期，2022年。[3] 世界卫生组织，“慕尼黑安全会议”。https://www.who.int/dg/speeches/detail/ munich-security-conference，2020年。访问日期：2022-12-09。[4] V. Buskens，社会网络与信任。荷兰多德雷赫特：Kluwer Academic Publishers，2002年。[5] V. Buskens and W. Raub，“嵌入式信任：控制与学习”，载于群体凝聚、信任与团结（E. J. Lawler and S. R. Thye 编），第167–202页，阿姆斯特丹：Elsevier，2002年。[6] C. Pechmann and S. J. Knight，“广告与同伴对青少年香烟消费信念和意图的联合影响：一项实验研究”，《消费者研究杂志》，第29卷，第1期，第5–19页，2002年。[7] S. Ozdemir, S. Zhang, S. Gupta, and G. Bebek，“信任和同伴影响对企业品牌—消费者关系及消费者忠诚度的影响”，《商业研究杂志》，第117卷，第791–805页，2020年。[8] Y. Wang and C. Yu，“社交商务中基于社会互动的消费者决策模型：口碑传播和观察学习的作用”，《国际信息管理杂志》，第37卷，第3期，第179–189页，2017年。

[9] D. C. Knill and A. Pouget，“贝叶斯大脑：不确定性在神经编码和计算中的作用”，《神经科学趋势》，第27卷，第12期，第712–719页，2004年。[10] J. K. Sebenius and Geanakoplos，“别打这个赌：信息不对称情况下的或有协议”，美国统计学会会刊，第78卷，第382期，第424–426页，1983年。[11] R. Parikh and P. Krasucki，“通信、共识与知识”，经济理论杂志，第52卷，第178–189页，1990年。[12] V. Bala and S. Goyal，“向邻居学习”，经济研究评论，第65卷，第3期，第595–621页，1998年。[13] J. Keppo, L. Smith, and D. Davydov，“最优选举时机：明智行事，或许能活得更久”，经济研究评论，第75卷，第597–628页，2008年。[14] M. Harel, E. Mossel, P. Strack, and O. Tamuz，“理性群体思维”，经济学季刊，第136卷，第1期，第621–668页，2021年。[15] S. Zhang and A. J. Yu，“便宜但聪明：多臂老虎机情境下的人类主动学习”，载于认知科学学会年会论文集，第35卷，Curran Associates, Inc.，2013年。[16] S. Zhang and A. J. Yu，“遗忘贝叶斯与短视规划：多臂老虎机情境下的人类学习与决策”，载于《神经信息处理系统进展》（C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger 编），第26卷，Curran Associates, Inc.，2013年。[17] P. Molavi, A. Tahbaz-Salehi, and A. Jadbabaie，“非贝叶斯社会学习理论”，Econometrica，第86卷，第2期，第445–490页，2018年。[18]W. Huang, P. Strack, and O. Tamuz，“网络中重复互动的学习”，2021年。arXiv:2112.14265v2。[19] M. H. DeGroot，“达成共识”，美国统计学会会刊，第69卷，第345期，第118–121页，1974年。[20] J. Correa, M. Mari, and A. Xia，“基于在线评论贝叶斯更新的动态定价”，NeurIPS经济政策机器学习研讨会，2020年。[21] W. Fu and A. Le Riche，“基于贝叶斯学习和技术选择的内生增长模型”，《数学社会科学》，第114卷，第58–71页，2021年。[22] A. Banerjee and D. Fudenberg，“口碑学习”，《博弈与经济行为》，第46卷，第1–22页，2004年。[23] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar，“社交网络中的贝叶斯学习”，经济研究评论，第78卷，第1201–1236页，2011年。[24] S. Bikhchandani, D. Hirschleifer, and I. Welch，“从他人行为中学习：从众、时尚潮流与信息级联”，《经济学展望杂志》，第12卷，第3期，第151–170页，1998年。[25] A. G. Bower, S. Garber, and J. C. Watson，“关于主体群体的学习以及信任与合作的演化”，《国际产业组织杂志》，第15卷，第165–190页，1996年。[26] V. Buskens，“三元组中的信任：退出、控制和学习的影响”，《博弈与经济行为》，第42卷，第235–252页，2003年。[27] V. Frey, V. Buskens, and W. Raub，“嵌入信任：关于网络嵌入性投资及其回报的博弈论模型”，《数学社会学杂志》，第39卷，第39–72页，2015年。[28] V. Buskens, W. Raub, and J. van der Veer，“三元组中的信任：一项实验研究”，《社会网络》，第32卷，第4期，第301–312页，2010年。[29] A. Kolb and E. Madsen，“疑云之下：秘密破坏下的信任动态”，经济研究评论，2022年6月。rdac034。[30]A. Kyprianou，《带应用的L´evy过程波动导论》。柏林：施普林格，2006年。[31] A. Kyprianou，“维纳–霍普夫分解。”载于：《定量金融百科全书》，R. Cont（编），2010年。[32] T. Cong and M. Sato，“受吸收屏障限制的不等步长一维随机游走”，《离散数学》，第40卷，第153–162页，1982年。

A. Appendix: Proofs of results

A.1. **引理1（保证退出）的证明**。

证明。 我们通过反证法证明这一点。假设对于所有 $t \in \mathbb{N}$，都有 $A_t = 1$。因此，

$$r\widehat{\vartheta}_t - c(1 - \widehat{\vartheta}_t) \geq 0, \quad \forall t \in \mathbb{N}. \tag{69}$$

代理人在第 t 轮通过以下方式估计可信度

$$\widehat{\vartheta}_t = \frac{\alpha + \widehat{S}_t}{\alpha + \beta + t}, \tag{70}$$

根据大数定律，我们知道

$$\lim_{t\to\infty} \frac{S_t}{t} = \vartheta \quad \text{(almost surely)}. \tag{71}$$

这也意味着

$$\lim_{t\to\infty} \widehat{\vartheta}_t = \vartheta < c/(r+c) \quad \text{(almost surely)}. \tag{72}$$

这意味着存在 t，使得

$$r\widehat{\vartheta}_t - c(1-\widehat{\vartheta}_t) < r\left(\frac{c}{c+r}\right) - c\left(1-\frac{c}{c+r}\right), \tag{73}$$

$$< \frac{rc}{c+r} - \frac{c(c+r)}{c+r} + \frac{c^2}{c+r} = \frac{0}{c+r}. \tag{74}$$

在 t 取此值时，交互的期望效用为负，这意味着代理人采取行动 $A_t = 0$，这与假设矛盾。□

A.2. **引理 2 的证明（收敛或退出）**。

证明。该代理人使用贝叶斯信念更新，从形状参数为 α 和 β 的贝塔分布先验信念开始。已知 $\widehat{S}_t = \sum_{t=1}^{t} X_t \mathbf{1}_{\{A_t=1\}}$，其中 X_t 是一个二元随机变量，以概率 ϑ 取值 1，以概率 $(1-\vartheta)$ 取值 0，其中 ϑ 是代理人试图估计的值，我们有时刻t 的信念分布等于

$$P_t(\theta) = \frac{\theta^{\widehat{S}_t+\alpha-1}(1-\theta)^{(t-\widehat{S}_t)+\beta-1}}{\int_0^1 y^{\widehat{S}_t+\alpha-1}(1-y)^{(t-\widehat{S}_t)+\beta-1}\mathrm{d}y}, \quad \theta \in [0,1]. \tag{75}$$

定义 τ 为第一个满足 $t \in \mathbb{N}$ 使得 $\int_0^1 \theta P_t(\theta)\,\mathrm{d}\theta < c/(c+r)$（可能为 ∞）的值，则对于所有$t < \tau$，估计的 ϑ 为

$$\widehat{\vartheta}_t = \frac{\alpha + \widehat{S}_t}{\alpha + \beta + t}. \tag{76}$$

于是有两种可能的情况：$\tau < \infty$ 和 $\tau = \infty$。在第一种情况下（其中 $\tau < \infty$），该代理人停止给予信任；根据关于 ϑ 的假设，即 $\vartheta \geq c/(c+r) = \theta_{\text{crit}}$，显然该代理人尚未收敛到 ϑ 的真实值。这是因为退出需要 $\vartheta \geq \theta_{\text{crit}}$ 这一事实，并且根据假设，我们选择了 $\vartheta > \theta_{\text{crit}}$。在第二种情况下，我们有 $\tau = \infty$。这里我们定义 S_t，从而消除对代理人行动的依赖：

$$S_t = \sum_{s=1}^{t} X_t, \quad \forall t \in \mathbb{N}, \tag{77}$$

并且我们注意到，当 $\tau = \infty$ 关系 $S_t = \widehat{S}_t$ 成立时，随之有 $\vartheta = (S_t + \alpha)/(\alpha + \beta + t)$。首先考察 S_t/t，根据大数定律，我们知道它需要

$$\lim_{t\to\infty} \frac{S_t}{t} = \vartheta \quad \text{(almost surely)}. \tag{78}$$

随后考察 ϑ_t 作为 $t \to \infty$，我们看到：

$$\lim_{t\to\infty} \widehat{\vartheta}_t = \lim_{t\to\infty} \frac{S_t + \alpha}{t + \alpha + \beta} = \vartheta \quad \text{(almost surely)}. \tag{79}$$

这表明 th 在代理人的估计值处，确实几乎必然收敛到真实的 ϑ 。 □

A.3. **引理 3 的证明（当 $c=1$ 且 $r\in\mathbb{N}$ 时的退出概率）**。

证明。 回顾假设 $\theta \geq c/(c+r)$ 以及定义 $\pi(u) = \mathbb{P}(\exists t:\ Z_t \geq u)$ 和 $\varrho = \mathbb{P}(\exists t:\ Z_t \geq 1)$。注意：(A) 根据强马尔可夫性质，随机游走到达水平 1 后的部分与此前发生的情况独立；(B) 在通往水平u的路径上，集合 $\{1,\ldots,u-1\}$ 中的任意水平都已被到达。因此我们得到“无记忆性”

$$\pi(u) = \varrho^u. \tag{80}$$

现在考虑从 0 开始的随机游走的第一步；它要么以概率 $(1-\theta)$ 立即到达第一个水平，要么下降到水平 $-r$，与水平 1 相距 $r+1$。利用（80），我们得到恒等式

$$F(\varrho) := \varrho = (1-\vartheta) + \vartheta\varrho^{r+1} =: G(\varrho). \tag{81}$$

我们可以证明，该方程在区间 $[0,1)$ 内恰好有一个解。注意，$F(\cdot)$ 是凸函数。显然，一个平凡解是 $\varrho = 1$ ，它构成凸函数与线性函数至多两个交点中的一个。因此，只需证明还存在另一个交点，且该交点位于 $[0,1)$ 内。确实如此，因为 $F'(1) = \theta(r+1)$，由于满足条件 $\theta > c/(c+r)$，代入 $c=1$ 并整理后，该值大于 $G'(1) = 1$。观察到 $F(0) = 1-\theta > 0 = G(0)$，这意味着在 $\varrho = 0$ 与 $\varrho = 1$ 之间的某处还存在另一个根。□

A.4. **引理 4 的证明（当 $c=1$ 且 $r\in\mathbb{N}$ 时的预期退出时间）**。

证明。 令 $\tau(u) = \inf\{t:\ Z_t = u\}$ 为随机游走 Z_t 到达 u 的首次穿越时间，并回忆一下，$\pi(u)$ 是该事件在有限时间内发生的概率（即，$\pi(u) := \mathbb{P}(\exists t:\ Z_t \geq u)$ ）。那么，我们感兴趣的是在 $\tau(u)$ 有限的条件下它的期望：

$$\mathbb{E}[\tau(u)\,|\,\tau(u)<\infty] = \frac{\mathbb{E}[\tau(u)\mathbf{1}\{\tau(u)<\infty\}]}{\pi(u)}. \tag{82}$$

注意这个关系

$$\mathbb{E}[z^{\tau(u)}\mathbf{1}\{\tau(u)<\infty\}] = \mathbb{E}[z^{\tau(1)}\mathbf{1}\{\tau(1)<\infty\}]^u \tag{83}$$

它与引理 3 的证明中完全相同的“无记忆论证”成立。我们定义

$$\varphi(z) := \mathbb{E}[z^{\tau(1)}\mathbf{1}\{\tau(1)<\infty\}], \tag{84}$$

根据关系式 (83)，这也允许我们得到一个自指表达式

$$\varphi(z) = z(1-\vartheta) + z\vartheta\,\varphi(z)^{r+1}. \tag{85}$$

最后请注意，我们可以通过求导并代入 $z=1$，再次得到所需的期望：

$$\mathbb{E}[\tau(u)\,|\,\tau(u)<\infty] = u\frac{\varphi'(1)}{\pi(1)}. \tag{86}$$

现在我们 对这些表达式进行处理，以得到最终结论：观察到 t

$$\varphi'(z) = (1-\vartheta) + \vartheta\varphi(z)^{r+1} + z\vartheta(r+1)\varphi(z)^r\varphi'(z), \tag{87}$$

替代 $z=1$ 并注意到 $\varphi(1)=\pi(1)=\varrho$，其中 ϱ 如式 (22) 所定义，我们得到

$$\varphi'(1)=\underbrace{(1-\vartheta)+\vartheta\varrho^{r+1}}_{=\varrho \text{ by (24)}}+\vartheta(r+1)\varrho^r\varphi'(1)=\frac{\varrho}{1-(r+1)\underbrace{\vartheta\varrho^r}_{=1-(1-\vartheta)\varrho^{-1}}} \tag{88}$$

$$=\frac{\varrho^2}{(1-\vartheta)(r+1)-r\varrho}\quad(\text{multiplying by }\frac{\varrho}{\varrho}).$$

将此乘以 u/ϱ，即可得到最终结果。 □

A.5. **引理 5 的证明（当 $r=1$，且 $c\in\mathbb{N}$ ）.**

证明。我们引入时间 G，它服从平移几何分布：$\mathbb{P}(G=t)=(1-f)^t f$，其中 $f\in(0,1)$。我们研究随机游走的行为，具体来说，是在几何触发的时间 G 之前到达 u 的概率：

$$p[f](u):=\mathbb{P}(\exists n\le G: Z_n\ge u), \tag{89}$$

以及对于某个 $n\le G$，到达 $Z_n=1$ 的概率：

$$\varrho[f]:=\mathbb{P}(\exists n\le G: Z_n\ge 1). \tag{90}$$

此外，我们将随机游走截至时间 n, $Z_n:=$ 的最小值定义为 $\{Z_0,Z_1,\dots,Z_n\}$，并将截至时间 n, $Z_n:=$ 所达到的最大值定义为 $\{Z_0,Z_1,\dots,Z_n\}$。由于维纳 - 霍普夫分解 [30, 31]，$\bar{Z}_G-Z_G$ 与 $\bar{Z}_G$ 相互独立，因此 $\mathbb{E}w^{Z_G}=\mathbb{E}w^{\bar{Z}_G}\mathbb{E}w^{Z_G-\bar{Z}_G}$. 它同样服从 Z_G，这一点可通过旋转随机游走 180°，并将 Z_G 视为新的、同分布过程的起点看出。因此，

$$\mathbb{E}w^{Z_G}=\mathbb{E}w^{\bar{Z}_G}\mathbb{E}w^{Z_G}. \tag{91}$$

因此，如果我们能够确定 $\mathbb{E}w^{Z_G}$ 和 $\mathbb{E}w^{Z_G}$，就找到了 $\mathbb{E}w^{\bar{Z}_G}$。为了得到 $\mathbb{E}w^{Z_G}$，令 x 表示该过程在时间区间 $t\in\{1,\dots,G\}$内向上跳跃的次数。因此，

$$\begin{aligned}\mathbb{E}w^{Z_G}&=\mathbb{E}\left[\sum_{x=0}^{G}\binom{G}{x}(1-\vartheta)^x\vartheta^{G-x}w^{cx-(G-x)}\right]\\&=\mathbb{E}\left[\sum_{x=0}^{G}\binom{G}{x}((1-\vartheta)w^c)^x\left(\frac{\vartheta}{w}\right)^{G-x}\right]=\mathbb{E}\left[(1-\vartheta)w^c+\vartheta w^{-1}\right]^G.\end{aligned} \tag{92}$$

回想一下，G 服从几何分布，对于 $|(1-f)((1-\theta)w^c+\theta w^{-1})|<1$,

$$\mathbb{E}w^{Z_G}=\sum_{i=0}^{\infty}(1-f)^i f((1-\vartheta)w^c+\vartheta w^{-1})^i=\frac{f}{1-(1-f)((1-\vartheta)w^c+\vartheta w^{-1})}. \tag{93}$$

为了求出 $\mathbb{E}w^{Z_G}$，我们注意到，最小值 Z_G 的分布与随机游走 B_G 最大值的相反数相同。该随机游走以概率 ϑ向上走一步（步长为 1），以概率 $1-\vartheta$向下走 $c\in\mathbb{N}$。这两种随机游走及其步长说明见表 4。随机游走 B_G 只是对文中分析的随机游走作了变量替换。

Walk	Step +	$\mathbb{P}(+)$	Step −	$\mathbb{P}(-)$
Z_G	c	$1-\vartheta$	1	ϑ
B_G	1	ϑ	c	$1-\vartheta$

TABLE 4. 两种随机游走。

在第3.2.1节中，于额外增加的偏移几何分布时间步 G 处，因此分析

极为相似。定义随机游走 B 在几何触发前至少上升 u 的概率，

$$p[f](u) := \mathbb{P}(\exists n \leq G : B_n \geq u).$$

S同理，定义在几何触发 b 之前到达 $u = 1$ 的概率 y

$$\varrho[f] := \mathbb{P}(\exists n \leq G : B_n \geq 1),$$

那么与之前一样，我们有无记忆性 $p[f](u) = (\varrho[f])^u$。显然，$\varrho[f]$ 也满足

$$\varrho[f] = (1-f)(\vartheta + (1-\vartheta)(\varrho[f])^{c+1}). \tag{94}$$

¯定义 $s[f](u) := \mathbb{P}(B_G = u)$ 为随机游走 B 在时刻 G 之前的最大值恰好为 u 的概率，并注意到

$$s[f](u) = p[f](u) - p[f](u+1) = (\varrho[f])^u - (\varrho[f])^{u+1} = (\varrho[f])^u(1-\varrho[f]), \tag{95}$$

我们得到了该最大值的表达式。根据 B_G 与 Z_G 之间的关系，我们也得到了 Z_G 的表达式，其中 $\varrho[f]$ 满足式（94）， —

$$\mathbb{P}(\underline{Z}_G = -u) = (\varrho[f])^u(1-\varrho[f]).$$

因此，

$$\mathbb{E}w^{\underline{Z}_G} = \sum_{u=0}^{\infty}(\varrho[f])^u(1-\varrho[f])w^{-u} = \frac{1-\varrho[f]}{1-\varrho[f]w^{-1}}. \tag{96}$$

结合表达式（93）和维纳-霍普夫分解（91），可得

$$\mathbb{E}w^{\bar{Z}_G} = \frac{f}{1-(1-f)((1-\vartheta)w^c + \vartheta w^{-1})} \times \frac{1-\varrho[f]w^{-1}}{1-\varrho[f]}. \tag{97}$$

通过考 ¯虑将其视为 $f \downarrow 0$，首先注意到 $\varrho[0] = 1$. 接着，定义 $\sigma(u) := \mathbb{P}(Z_\infty = u$),

$$\begin{aligned}
\zeta(w) &:= \sum_{u=0}^{\infty}\sigma(u)w^u = \lim_{f\downarrow 0}\mathbb{E}w^{\bar{Z}_G} \\
&= \lim_{f\downarrow 0}\frac{f}{1-(1-f)((1-\vartheta)w^c + \vartheta w^{-1})} \times \frac{1-\varrho[f]w^{-1}}{1-\varrho[f]} \\
&= \lim_{f\downarrow 0}\frac{1-\varrho[f]w^{-1}}{1-(1-f)((1-\vartheta)w^c + \vartheta w^{-1})} \times \frac{f}{1-\varrho[f]} \\
&= \frac{1-w^{-1}}{1-((1-\vartheta)w^c + \vartheta w^{-1})} \times \lim_{f\downarrow 0}\frac{f}{1-\varrho[f]}
\end{aligned} \tag{98}$$

应用洛必达法则，可得

$$\zeta(w) = \frac{1-w^{-1}}{1-((1-\vartheta)w^c + \vartheta w^{-1})} \times \frac{1}{-\gamma}, \quad \text{with} \quad \gamma := \lim_{f\downarrow 0}\frac{\mathrm{d}}{\mathrm{d}f}\varrho[f]. \tag{99}$$

通过对 (94) 两边关于 f 求导并求解 γ，很容易验证 $1/\gamma = (c+1)(1-\vartheta) - 1$，由此可得

$$\zeta(w) = \frac{(1-w^{-1})((c+1)(1-\vartheta)-1)}{((1-\vartheta)w^c + \vartheta w^{-1}) - 1} = \frac{(w-1)(1-(1-\vartheta)(c+1))}{w - \vartheta - (1-\vartheta)w^{c+1}}. \tag{100}$$

但我们关注的是 $\pi(u)$，因此令 $\xi(w) = \sum_{u=0}^{\infty} \pi(u) w^u$，用$\sigma(u)$ 可将其表示为

$$\begin{aligned}\xi(w) &= \sum_{u=0}^{\infty}\sum_{v=u}^{\infty} \sigma(v) w^u = \sum_{v=0}^{\infty}\sum_{u=0}^{v} \sigma(v) w^u \\ &= \sum_{v=0}^{\infty} \sigma(v) \left(\frac{1-w^{v+1}}{1-w}\right) = \frac{\sum_{v=0}^{\infty}\sigma(v) - w\sum_{v=0}^{\infty}\sigma(v) w^v}{1-w},\end{aligned} \tag{101}$$

最终得到这个有用的表达式

$$\xi(w) = \frac{1 - w\zeta(w)}{1-w}, \tag{102}$$

借此我们可以通过关系求出 $\pi(u)$

$$\pi(u) = \frac{1}{u}\frac{\mathrm{d}^u}{\mathrm{d}w^u}\xi(w)\bigg|_{w=0} = \frac{\xi^{(u)}(0)}{u!}. \tag{103}$$

将 $\zeta(w)$ 的表达式代入即可得到结果。 □

注记 7. Cong 和 Sato [32] 使用组合论证，求出在特定时刻首次到达吸收屏障的概率 t，适用于以下两种情形：r 或 c 为 1，

$$p_t(u) = \mathbb{P}(\inf\{n \in \mathbb{N} : Z_n \geq u\} = t). \tag{104}$$

对首次到达时间 t 求和，进而可以得到我们在此给出的概率：$\sum_{t=1}^{\infty} p_t(u) = \pi(u)$。请注意，这种方法本质上是近似的，因为必须以某种方式截断对 t 的求和；而我们的方法是精确的。我们没有使用他们的结果，原因之一是其相关计算十分复杂，此外，该结果也无法给出预期退出时间。

A.6. **引理 6 的证明（当 $r = 1$ 且 $c \in \mathbb{N}$ 时的预期退出时间）**。

证明。定义

$$\varphi(z,u) := \mathbb{E}[z^{\tau(u)}\mathbf{1}\{\tau(u) < \infty\}], \quad \Phi(w,z) := \sum_{u=1}^{\infty} w^u \varphi(z,u), \quad w \in (-1,1), \tag{105}$$

使得

$$\mathbb{E}[\tau(u) \,|\, \tau(u) < \infty] = \frac{\varphi'(1,u)}{\varphi(1,u)}, \quad \varphi(z,u) = \frac{\partial^u}{\partial w^u}\frac{\Phi(w,z)}{u!}\bigg|_{w=0} \tag{106}$$

过程动力学揭示了递归关系

$$\varphi(z,u) = \begin{cases} (1-\vartheta)z\,\varphi(z,u-c) + \vartheta z\,\varphi(z,u+1) & \text{if } u > c, \\ (1-\vartheta)z + \vartheta z\,\varphi(z,u+1) & \text{otherwise.} \end{cases} \tag{107}$$

代入 将其代入 $\Phi(w,\ z)$ 的定义，经过一些简单的改写后，我们得到 g,

$$\Phi(w,z) = \underbrace{\sum_{u=1}^{c} w^u(1-\vartheta)z}_{(a)} + \underbrace{\sum_{u=1}^{\infty} w^u \vartheta z\,\varphi(z,u+1)}_{(b)} + \underbrace{\sum_{u=c+1}^{\infty} w^u(1-\vartheta)z\,\varphi(z,u-c)}_{(c)}. \tag{108}$$

分别考察（a）、（b）和（c），我们得到

$$(a) = (1-\vartheta)z\sum_{u=1}^{c} w^u = (1-\vartheta)z\left(\frac{w - w^{c+1}}{1-w}\right) \tag{109}$$

应用有限等比级数后，

$$(b) = \vartheta z \frac{1}{w} \sum_{u=1}^{\infty} w^{u+1} \varphi(z, u+1) = \frac{\vartheta z}{w} \Phi(w, z) - \vartheta z \, \varphi(z, 1) \tag{110}$$

根据 $\Phi(w, z)$ 的定义，并且

$$(c) = w^c (1-\vartheta) z \sum_{u=c+1}^{\infty} w^{u-c} \varphi(z, u-c) = w^c (1-\vartheta) z \, \Phi(w, z) \tag{111}$$

将求和指标平移 $c+1$，并再次应用 $\Phi(w, z)$ 的定义。通过合并并重新整理（a）、（b）和（c），我们可以分离出 $\Phi(w, z)$，从而得到

$$\Phi(w, z) = \frac{(1-\vartheta) z \left(\frac{w - w^{c+1}}{1-w} \right) - \vartheta z \varphi(z, 1)}{1 - (\vartheta z)/w - w^c z (1-\vartheta)}. \tag{112}$$

接下来需要确定对于每个 $z \in (-1, 1)$，$\varphi(z,1)$ 的值。关键原则是，由于它们的比值是有限数，(112) 中分母的任何根也必然是分子的根。事实上，假设存在唯一的 $w(z) \in (-1, 1)$，使分母取得一个根，那么这将意味着

$$\varphi(z, 1) = \frac{1-\vartheta}{\vartheta} z \left(\frac{w(z) - w(z)^{c+1}}{1 - w(z)} \right) = \frac{(1-\vartheta) z w(z) + \vartheta z - w(z)}{1 - w(z)}. \tag{113}$$

这意味着，只要我们能够证明：对于任意 $z \in (-1,1)$，(112) 中的分母在（-1,1）中存在唯一的根 $w(z)$。可以直接看出，该根满足

$$F(w) := w^{c+1} z (1-\vartheta) + z\vartheta = w =: G(w). \tag{114}$$

注意到，对于 $z = 0$，我们显然有唯一根 $w = 0$.

子情形 $z > 0$。注意到 $F(0) = z\vartheta > 0 = G(0)$ 且 $F(1) = z < 1 = G(1)$，因此 $F(w) - G(w)$ 对于 $w \in (0, 1)$ 变号奇数次。但由于 $F(\cdot)$ 是凸的，而 $G(\cdot)$ 是线性的，函数 $F(\cdot) - G(\cdot)$ 至多变号两次。结合以上结果，我们得出唯一根 $w(z) \in (-1, 1)$ 存在（实际上该根位于 0 和 1 之间）。

子情形 $z < 0$。c 为奇数的情形与子情形 $z > 0$ 类似：$F(0) = z\vartheta < 0 = G(0)$ 且 $F(-1) = z > -1 = G(-1)$。根据 $F(\cdot)$ 的凹性，可知唯一根 $w(z) \in (-1, 1)$ 存在（该根位于 -1 和 0 之间）。当 c 为偶数时，我们仍有 $F(0) = z\vartheta < 0 = G(0)$，但现在应注意到 $F(-1) = z(2\vartheta - 1)$；对于 $z \in (-1, 0)$ 和 $\vartheta \in (0, 1)$，其值大于 $-1 = G(-1)$（取 $\vartheta = 1$ 和 $z = -1$）。现在，由于当 $w \in (-1, 0)$ 时 $F(w)$ 递减而 $G(w)$ 递增，可得唯一根 $w(z) \in (-1, 1)$ 存在（该根位于 -1 和 0 之间）。□

B. Appendix: Single agent case r, $c \in \mathbb{N}$

在这种情况下，无法得出退出概率或以此为条件的预期退出时间的精确公式；回想一下，在另外两种情况下，我们利用了随机游走无论向上还是向下，步长均为 1 这一事实。下面简要概述一种能够准确近似退出概率的数值方案。对于所有 $u > c$，关系式为：

$$\pi(u) = \vartheta \pi(u + r) + (1-\vartheta) \pi(u - c). \tag{115}$$

对于足够大的 $u > u_0$，我们通过设定 $\widehat{\pi}(u) = \eta\gamma^u$ （$\eta \geq 0$ 和 $\gamma \in (0, 1)$）来补充这一近似。结合 (115)，我们知道 γ 满足 $1 = (1-\vartheta)\gamma^{-c} + \vartheta\,\gamma^{r}$。使用来自 $u \geq u_0$ 的 $\pi(u)$ 近似，可得：

$$\widehat{\pi}(u) = \begin{cases} \eta\gamma^u & \text{if } u \geq u_0 \\ (1-\vartheta)\widehat{\pi}(u-c) + \vartheta\widehat{\pi}(u+r) & \text{if } c < u < u_0 \\ (1-\vartheta) + \vartheta\widehat{\pi}(u+r) & \text{if } 0 < u \leq c, \end{cases} \tag{116}$$

这可用于构造一个能够通过数值方法求解的方程组。我们提出用该方程组（116）在适当的 $u = r\alpha - c\beta + 1$ 下的解，来近似吸收概率（以及随后发生的信任终止）。u_0 的值应足够大，以获得良好的近似，同时又应足够小，以使其计算保持在可控范围内。

我们概述一种类似的系统中的期望时间近似方案。该方案的核心是在足够高的 u 下，用在不采取任何远离吸收的步骤的情况下直接向吸收移动所需的轮次数，来近似预期退出时间：$\lceil u/c \rceil$。就 $\varphi(z, u)$ 而言，可以得到以下方程组：

$$\widehat{\varphi}(z,u) = \begin{cases} (1-\vartheta)^{\lceil u/c \rceil} z^{\lceil u/c \rceil} & \text{if } u \geq u_0 \\ (1-\vartheta)z\,\varphi(z, u-c) + \vartheta z\,\varphi(z, u+r) & \text{if } c < u < u_0 \\ (1-\vartheta)z + \vartheta z\,\varphi(z, u+r) & \text{if } 0 < u \leq c. \end{cases} \tag{117}$$

这定义了一个类似于（116）的方程组，可以通过数值方法求解。由此可得到 $\varphi(z, u_{\text{crit}})$ 的一个近似值，进而可通过在 u_{crit} 处求导来得到 $\mathbb{E}[\tau(u_{\text{crit}}) \mid \tau(u_{\text{crit}}) < \infty]$ 的近似值：$\varphi'(1, u_{\text{crit}}) / \varphi(1, u_{\text{crit}})$。注意其与（40）的相似之处。

C. Appendix: An illustration of how the OA model can beat the OR model

在本附录中，我们详细探讨一个特殊结果：在参数设定$c = 3$、 $r = 2$和 $u^* = 2$ 下，当可信度设定为 $\vartheta = 0.65$，0。66时，OA模型的退出概率低于OR模型。

我们考虑一对代理人，所处情形与该设定完全相同，唯一区别是（$c = 3$、 $r = 2$、$u^* = 2$ 和 $\vartheta = 0.65$）。此外，我们假设机构前四次受到信任（由任一代理人给予）时，只有一次信任得到了回报。

在OR模型中，这意味着交互最多只能进行两个轮次。第一轮中，两位代理人的信任都可能遭到滥用，在这种情况下，他们会在第1轮退出；或者第一轮中的信任得到一次回报，但第二轮中信任遭到两次滥用，因此代理人不会在第3轮给予信任。

在OA模型中，无论第一轮次获得的奖励如何，参与其中的主体都会进行两轮次交互。在第2轮次开始时给予信任的行为不会向邻居传递任何信息。然而，在第3轮次开始时，观察到两次滥用信任行为的主体（标记为1）不再给予信任。这向邻近主体（标记为2）发出信号，表明滥用信任的行为确实发生了两次。在第3轮次给予信任时（在观察该行为的结果之前），主体2能够获取相关信息，使其估计值低于临界值。然而，该主体已经给予了信任，因此仍会看到这次交互的结果。这一结果至关重要，因为如果信任得到履行（随着 $\vartheta = 0.65 > 0.5$而更有可能），主体2将继续维持这段关系；而如果信任遭到滥用，他们就会退出。

D. Appendix: Tables of simulated results

在下表中，我们展示了在所考虑的三种机制下退出的概率以及预期退出时间。

TABLE 5. 单智能体模型（Sol）、可观察行动模型（OA）和可观察奖励模型（OR）中的退出概率和预期退出时间。参数 α 和 β 的设置满足 $u^* = 1$

(A)退出概率

$c = r = 1$													$\alpha = \beta = 2$
ϑ	0.18	0.24	0.30	0.36	0.42	0.45	0.55	0.60	0.66	0.72	0.78	0.84	0.90
Sol	1.0	1.0	1.0	1.0	1.0	1.000	0.808	0.669	0.518	0.390	0.288	0.198	0.119
OA	1.0	1.0	1.0	1.0	1.0	0.999	0.797	0.649	0.493	0.370	0.272	0.186	0.113
OR	1.0	1.0	1.0	1.0	1.0	1.000	0.648	0.429	0.248	0.138	0.072	0.030	0.010
$c < r$:													
$c = 2, r = 3$													$\alpha = 3, \beta = 4$
ϑ	0.18	0.24	0.3	0.35	0.45	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	0.832	0.755	0.611	0.495	0.394	0.311	0.241	0.170	0.108
OA	1.0	1.0	1.0	1.0	0.825	0.734	0.592	0.483	0.382	0.296	0.230	0.165	0.102
OR	1.0	1.0	1.0	1.0	0.745	0.621	0.424	0.292	0.190	0.120	0.066	0.030	0.010
$c = 1, r = 2$													$\alpha = 2, \beta = 4$
ϑ	0.18	0.24	0.28	0.38	0.42	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	0.869	0.777	0.658	0.551	0.459	0.377	0.302	0.234	0.169	0.108
OA	1.0	1.0	1.0	0.866	0.766	0.641	0.538	0.447	0.366	0.289	0.225	0.164	0.102
OR	1.0	1.0	1.0	0.787	0.644	0.478	0.336	0.243	0.162	0.105	0.060	0.028	0.010
$c > r$:													
$c = 3, r = 2$													$\alpha = 4, \beta = 2$
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.65	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.743	0.703	0.487	0.330	0.215	0.124
OA	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.721	0.682	0.475	0.318	0.205	0.120
OR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.657	0.600	0.333	0.168	0.066	0.024
$c = 2, r = 1$													$\alpha = 5, \beta = 2$
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.6	0.62	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.999	0.682	0.438	0.27	0.142
OA	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.000	0.667	0.433	0.26	0.140
OR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.000	0.622	0.352	0.17	0.068

(B)预期退出时间

$c = r = 1$													$\alpha = 2, \beta = 2$
ϑ	0.18	0.24	0.30	0.36	0.42	0.45	0.55	0.60	0.66	0.72	0.78	0.84	0.90
Sol	1.550	1.928	2.546	3.640	6.702	10.154	9.236	4.841	3.159	2.184	1.713	1.510	1.252
OA	1.352	1.597	2.037	2.740	4.718	7.710	7.116	3.891	2.480	1.883	1.478	1.237	1.155
OR	1.562	1.905	2.482	3.701	6.852	11.253	9.882	5.529	3.193	2.177	1.710	1.426	1.095
$c < r$:													
$c = 2, r = 3$													$\alpha = 3, \beta = 4$
ϑ	0.18	0.24	0.3	0.35	0.45	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.897	2.711	4.408	9.158	8.613	5.552	3.033	2.136	1.619	1.388	1.262	1.089	1.007
OA	1.690	2.231	3.584	7.342	7.082	4.439	2.702	2.057	1.532	1.308	1.184	1.121	1.039
OR	1.668	2.197	3.329	7.241	7.834	4.925	2.882	2.096	1.855	1.583	1.397	1.339	1.095
$c = 1, r = 2$													$\alpha = 2, \beta = 4$
ϑ	0.18	0.24	0.28	0.38	0.42	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	2.130	3.558	6.052	6.755	4.045	2.690	1.960	1.541	1.352	1.243	1.112	1.053	1.007
OA	1.898	2.904	5.278	6.204	3.804	2.277	1.770	1.463	1.282	1.192	1.102	1.084	1.039
OR	1.874	2.975	4.855	6.630	3.898	2.506	1.932	1.724	1.517	1.352	1.281	1.218	1.095
$c > r$:													
$c = 3, r = 2$													$\alpha = 4, \beta = 2$
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.65	0.66	0.72	0.78	0.84	0.9
Sol	1.514	1.824	2.250	2.831	3.856	5.650	11.644	12.538	10.590	5.090	2.762	1.988	1.497
OA	1.328	1.531	1.833	2.242	2.941	4.552	10.004	9.398	7.988	4.010	2.388	1.750	1.418
OR	1.470	1.706	2.014	2.476	3.161	4.583	9.070	10.020	8.075	4.647	3.172	2.500	2.250
$c = 2, r = 1$													$\alpha = 5, \beta = 2$
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.6	0.62	0.72	0.78	0.84	0.9
Sol	1.444	1.692	2.004	2.414	3.034	3.958	5.837	10.977	15.353	12.359	5.390	3.174	1.961
OA	1.270	1.418	1.627	1.891	2.332	3.072	4.684	9.232	13.028	9.437	4.308	2.448	1.734
OR	1.360	1.522	1.683	1.954	2.350	3.007	4.348	7.880	11.248	9.034	4.131	3.009	2.336

TABLE 6. 单智能体模型（Sol）、可观察行动模型（OA）和可观察奖励模型（OR）中的退出概率和预期退出时间。参数 α 和 β 设置为使得 $u^* = 2$

(A)退出概率

$c = r = 1$												$\alpha = 3, \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.45	0.55	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	1.0	1.000	0.665	0.450	0.267	0.159	0.086	0.038	0.012
OA	1.0	1.0	1.0	1.0	1.0	0.999	0.647	0.429	0.252	0.141	0.078	0.034	0.012
OR	1.0	1.0	1.0	1.0	1.0	1.000	0.648	0.429	0.248	0.138	0.072	0.030	0.010
$c < r$:													
$c = 2, r = 3$												$\alpha = 3, \beta = 3$	
ϑ	0.18	0.24	0.3	0.35	0.45	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	0.700	0.574	0.376	0.249	0.160	0.101	0.060	0.028	0.011
OA	1.0	1.0	1.0	1.0	0.664	0.532	0.356	0.228	0.146	0.089	0.052	0.025	0.010
OR	1.0	1.0	1.0	1.0	0.659	0.526	0.328	0.204	0.124	0.082	0.048	0.022	0.010
$c = 1, r = 2$												$\alpha = 2, \beta = 3$	
ϑ	0.18	0.24	0.28	0.38	0.42	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	0.760	0.604	0.434	0.303	0.214	0.145	0.096	0.058	0.028	0.011
OA	1.0	1.0	1.0	0.747	0.589	0.418	0.297	0.204	0.136	0.086	0.051	0.025	0.010
OR	1.0	1.0	1.0	0.724	0.559	0.386	0.258	0.176	0.115	0.080	0.048	0.022	0.010
$c > r$:													
$c = 3, r = 2$												$\alpha = 7, \beta = 3$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.65	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.572	0.508	0.241	0.115	0.047	0.014
OA	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.526	0.463	0.224	0.104	0.039	0.013
OR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.542	0.481	0.208	0.088	0.032	0.010
$c = 2, r = 1$												$\alpha = 7, \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.6	0.62	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.998	0.482	0.202	0.078	0.022
OA	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.999	0.466	0.197	0.072	0.021
OR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.000	0.441	0.164	0.048	0.013

(B)预期退出时间

$c = r = 1$												$\alpha = 3, \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.45	0.55	0.6	0.66	0.72	0.78	0.84	0.9
Sol	3.092	3.818	4.970	7.224	12.957	20.029	18.126	9.550	6.213	4.431	3.455	3.000	2.130
OA	2.615	3.077	3.863	5.228	9.552	15.800	14.165	7.130	4.683	3.664	3.172	2.597	2.362
OR	1.562	1.905	2.482	3.701	6.852	11.253	9.882	5.529	3.193	2.177	1.710	1.426	1.095
$c < r$:													
$c = 2, r = 3$												$\alpha = 3, \beta = 3$	
ϑ	0.18	0.24	0.3	0.35	0.45	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	3.728	5.220	8.389	17.154	16.696	11.193	6.267	4.447	3.353	2.737	2.432	2.126	2.000
OA	3.060	4.100	6.736	14.250	13.397	7.572	5.232	3.762	2.888	2.511	2.203	2.148	2.026
OR	2.146	3.056	4.791	10.426	10.130	6.521	3.615	2.262	1.634	1.323	1.146	1.068	1.000
$c = 1, r = 2$												$\alpha = 2, \beta = 3$	
ϑ	0.18	0.24	0.28	0.38	0.42	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	4.173	7.032	11.962	14.484	8.156	5.432	3.862	3.085	2.767	2.584	2.258	2.082	2.000
OA	3.364	5.240	9.371	11.275	6.624	4.143	3.223	2.668	2.415	2.311	2.138	2.091	2.026
OR	2.384	3.978	6.820	8.322	5.202	2.833	1.959	1.586	1.300	1.189	1.146	1.068	1.000
$c > r$:													
$c = 3, r = 2$												$\alpha = 7, \beta = 3$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.65	0.66	0.72	0.78	0.84	0.9
Sol	2.992	3.560	4.337	5.489	7.374	10.940	22.088	25.305	20.474	9.427	5.633	4.096	3.000
OA	2.559	2.906	3.429	4.159	5.572	8.398	17.574	16.967	14.051	7.563	4.701	3.219	2.755
OR	1.551	1.869	2.298	3.072	4.174	6.394	12.963	13.472	11.554	5.665	2.780	1.656	1.095
$c = 2, r = 1$												$\alpha = 7, \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.6	0.62	0.72	0.78	0.84	0.9
Sol	2.851	3.288	3.874	4.686	5.866	7.657	11.462	21.216	30.238	25.010	10.048	6.758	4.529
OA	2.467	2.712	3.087	3.552	4.325	5.639	8.380	16.212	23.482	16.732	7.119	4.745	3.602
OR	1.520	1.794	2.142	2.666	3.360	4.418	6.550	12.495	17.276	13.907	5.960	3.320	1.731

TABLE 7. 单智能体模型（Sol）、可观察行动模型（OA）和可观察奖励模型（OR）中的退出概率和预期退出时间。参数 α 和 β 设定为使得 $u^* = 3$

(A)预期退出时间

$c = r = 1$												$\alpha = 4, \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.45	0.55	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	1.0	1.000	0.544	0.304	0.138	0.063	0.025	0.006	0.001
OA	1.0	1.0	1.0	1.0	1.0	0.999	0.526	0.293	0.139	0.068	0.024	0.007 0.001	
OR	1.0	1.0	1.0	1.0	1.0	1.000	0.441	0.194	0.062	0.019	0.006	0.000	0.000
$c < r$:													
$c = 2, r = 3$												$\alpha = 5, \beta = 5$	
ϑ	0.18	0.24	0.3	0.35	0.45	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	0.999	0.594	0.436	0.238	0.128	0.066	0.033	0.013	0.004	0.001
OA	1.0	1.0	1.0	1.000	0.564	0.419	0.232	0.124	0.067 0.035		0.014	0.004	0.001
OR	1.0	1.0	1.0	1.000 0.546 0.386			0.186 0.084		0.030 0.012		0.005	0.000 0.000	
$c = 1, r = 2$												$\alpha = 2, \beta = 2$	
ϑ	0.18	0.24	0.28	0.38	0.42	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	0.662	0.470	0.288	0.170	0.100	0.055	0.028	0.013	0.004	0.001
OA	1.0	1.0	1.0	0.622	0.441	0.264	0.152	0.092	0.051	0.028	0.012	0.004	0.001
OR	1.0	1.0	1.0	0.610	0.406	0.216	0.114	0.051	0.017 0.006		0.003	0.000	0.000
$c < r$:													
$c = 3, r = 2$												$\alpha = 7, \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.65	0.66	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.439	0.367	0.129	0.040	0.011	0.001
OA	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.396	0.317	0.116	0.036	0.008	0.001
OR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.404	0.334	0.100	0.025	0.004	0.000
$c = 2, r = 1$												$\alpha = 9, \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.6	0.62	0.72	0.78	0.84	0.9
Sol	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.996	0.344	0.096	0.022	0.002
OA	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.999	0.320	0.079	0.015	0.002
OR	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.999	0.315	0.076	0.011	0.000

(B)预期退出时间

$c = r = 1$												$\alpha = 4, \ \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.45	0.55	0.6	0.66	0.72	0.78	0.84	0.9
Sol	3.966	4.680	5.788	7.518	10.987	19.082	29.922	28.213	15.061	9.271	6.518	5.260	4.760
OA	3.463 3.834		4.468	5.503	7.552 13.513 22.291 19.470				10.387	6.623	5.273	4.385 3.815	
OR	2.646 3.140		3.893	5.038	7.396	13.353	21.750	20.367 11.100		6.712	4.921	3.454 4.000	
$c < r$:													
$c = 2, r = 3$												$\alpha = 5, \beta = 5$	
ϑ	0.18	0.24	0.3	0.35	0.45	0.48	0.54	0.6	0.66	0.72	0.78	0.84	0.9
Sol	4.487 5.678		7.859	12.435	24.902	26.330	16.966	9.608	6.672	5.149	4.614	3.623	3.429
OA	3.680 4.426		5.845	9.346 20.049 18.274 11.024				7.528	5.453	4.598	3.993	3.536 3.389	
OR	2.724	3.393	4.652	7.328	15.510	15.062 10.206		5.879	4.389	3.881	3.522	3.100	4.000
$c = 1, r = 2$												$\alpha = 2, \beta = 2$	
ϑ	0.12	0.18	0.24	0.28	0.38	0.42	0.48	0.54	0.6	0.66	0.72	0.78	0.84
Sol	4.733	6.418	10.601	18.015	21.900	12.759	8.160	5.753	4.537	4.059	3.868	3.471	3.429
OA	3.911	5.140	8.152	14.868	16.413	10.208	6.396	4.823	4.194	3.628	3.460	3.229 3.000	
OR	2.912 3.972		6.327 10.711 13.418			8.477	4.984	3.833	3.157	2.794	2.546	2.500	Null
$c > r$:													
$c = 3, \ r = 2$												$\alpha = 7, \ \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.65	0.66	0.72	0.78	0.84	0.9
Sol	3.911	4.511	5.332	6.422	8.088	10.913	16.426	32.571	36.403	29.336	13.765	8.543	7.140
OA	3.457 3.799		4.323	5.068	6.141	8.104 12.119 24.871 25.874 20.474 10.281						7.240	4.645
OR	2.570 2.910		3.344	3.907	4.867	6.274	9.379 18.112 19.234 17.288				8.740	4.880 4.000	
$c = 2, r = 1$												$\alpha = 9, \beta = 2$	
ϑ	0.18	0.24	0.3	0.36	0.42	0.48	0.54	0.6	0.62	0.72	0.78	0.84	0.9
Sol	3.815	4.305	4.963	5.782	6.950	8.716	11.345	16.756	31.361	43.763	37.012	15.414	9.261
OA	3.446 3.740		4.151	4.683	5.392	6.558	8.474 12.322 23.541 33.944 25.088 10.330						7.312
OR	2.442 2.682		2.984	3.358	4.006	4.910	6.421	9.459	17.515 24.158 20.682			9.183	4.500