arXiv:2601.08642v1 [cs.GT] 13 Jan 2026

城市中的博弈：改善城市邻里博弈中的均衡

Martin Gairing* Adrian Vetta† 赵展展‡

2026年1月14日

摘要

当个体会根据本地情况采取策略性行动时，城市应如何投资以提升社会福利？我们使用谢林有界邻域模型的博弈论版本来建模这一问题，其中主体根据反映街区人口的凹且非单调效用函数选择街区。尽管看似简单的改进可能使结果恶化——类似于布雷斯悖论——我们证明，精心设计的小规模投资能够可靠地使个体激励与社会目标保持一致。具体而言，以至多为 $0.81\epsilon^2 \cdot$ opt的总成本调整效用，可以保证由此产生的每个纳什均衡都能实现至少为 $\epsilon \cdot$opt的社会福利，其中opt为最优社会福利。我们的结果形式化地说明，定向干预如何将负效应转化为正收益，并为策略性城市规划和去中心化集体行为提供了新的见解。

1 引言

社会应当如何、何时以及在哪里投入资源，才能最有效地提升市民的效用？我们在城市规划的背景下探讨这一问题，重点关注街区改善。举例来说，假设某个大都市拥有 1 亿美元的发展预算，可用于道路工程、游乐场、公园等项目。古典经济学预测，这项投资可能带来的社会效益从零效益（即浪费资金时，应将 0 美元返还给纳税人）到全部效益（即 1 亿美元；考虑到适度的乘数效应，效益可能略高）不等。然而，令人惊讶的是，这大大低估了这项投资可能产生的结果范围。原因在于，这种预测忽视了城市规划者的宏观动机与市民个体的微观动机之间的区别。具体而言，个人激励可能促使个人对投资作出反应，从而对其他市民的效用产生正面或负面影响。在我们的研究设定中，我们关注的相应微观行为是迁居至其他街区，并使用谢林隔离模型的一种标准变体来量化其影响 [29]。

谢林 [29] 探讨了一个动态模型：两类参与者占据图中的节点，并对其所在街区内智能体类型的构成具有偏好。他的研究表明，即使个体只是稍微偏好与自己同类的邻居，也可能导致极端

*利物浦大学：gairing@liverpool.ac.uk
†麦吉尔大学：adrian.vetta@mcgill.ca
‡香港中文大学（深圳）：zhanzhanzhao@cuhk.edu.cn

隔离，即便个体也愿意接受一定程度的融合。谢林模型适用于一般图，同时也探讨了有界邻域模型：城市被划分为街区，参与者根据街区的整体人口构成决定进入或离开某个街区。有界邻域模型得到了更广泛的研究（例如参见 [12, 13]）；特别是，Grauwin 等人 [13] 对谢林模型进行了简化，提出一种只有一种参与者类型的模型，其中主体效用取决于其所在街区的主体密度。作者展示了如何将其单人口模型的结果直接扩展到谢林的双人口模型。其主要贡献是基于物理学中使用的自由能函数，对稳态进行解析预测，从而无需再进行以往必需的数值模拟。

在本文中，我们研究了与 Grauwin 等人相同的模型 [13] ，但采用博弈论视角。我们的研究展示了如何引导宏观投资，以优化均衡状态下的社会福利；也就是说，我们使微观动机与社会福利目标保持一致。微观动机所带来的负面影响有一个引人注目的例子，即布雷斯悖论 [8] ：在城市中修建一条新道路，竟可能导致所有人的出行时间都增加！这是因为新道路可能会在道路网络中造成瓶颈。布雷斯悖论在实践中已屡见不鲜。事实上，从理论角度看，这种现象可能“十分常见” [30]。反过来，通常可以看到一些可视为撤资的举措（例如封闭道路、采取交通缓和措施、征收道路收费等）反而会提高所有人的交通速度。

另一方面，微观行为也能产生不成比例的积极影响。一个著名的例子是毕尔巴鄂效应 [22, 23]。据估算，毕尔巴鄂古根海姆博物馆每年为当地经济带来的经济收益达 6.5 亿美元。因此，与旅游业和经济振兴相关的微观行为所创造的收益，远远超过了博物馆 1 亿美元的建造成本。

因此，社区投资后可实现的结果范围从超负向到超正向。我们将在正式定义城市规划模型后，于第 2 节展示这两种情况的示例。

本文旨在确定何时以及如何实现超正向投资回报。我们的结论相当令人惊讶：要么当前社区规划所提供的社会福利已与当时可实现的最优社会福利相当，要么存在能够带来超正向回报的针对性社区投资。

为形式化并量化这一结论，我们采用谢林有界邻域模型的一个广义版本 [29]。我们在第 2 节介绍该模型，并通过示例说明将用于分析该模型的博弈论工具。随后，我们在第 3 节讨论相关文献。在第 5 节，我们证明主要结果——定理1：如果当前邻域均衡的福利与可实现的最优社会福利相比偏低，那么超正向投资回报总是存在。在第6节，我们证明一个互补结果——定理2：定理1所保证的投资回报基本上已是最佳可能结果，也就是说，其保证范围在一个较小的乘法因子内是紧的。最后，我们在第 7 节作出总结。

2 城市规划的博弈论模型

对于我们的城市规划模型，我们采用谢林有界邻域模型的一个推广版本 [29]。大都市中有 n 个街区，其中每个街区 i的最大人口容量为 u_i ，人口为 $x_i \in [0, u_i]$。因此，大都市的总人口为 $N = \sum_{i=1}^{n} x_i$。为使表述更简洁，我们采用人口的非原子模型

，在该模型中，人口由连续统中的无限多个个体构成，每个个体对人口的贡献都无穷小。鉴于人口通常规模庞大，且每个个体对效用的影响都微乎其微，这种设定很好地近似了我们的情形。此外，我们强调，所有结果和证明同样适用于个体数量有限的原子版本（作相应调整即可）。

这里有两点需要说明。首先，谢林的原始模型 [29] 中，人口由两个种族构成，在街区 i 中，两者的人口密度分别为 d_i 和 $1-d_i$。就分析而言，这与我们的设定等价；为说明这一点，令人口 $x_i = d_i \cdot u_i$ 对应第一种族的人口，剩余容量 $u_i - x_i = (1-d_i) \cdot u_i$ 对应第二种族的人口。其次，我们的研究动机适用于大都市中的街区尺度。但我们的结果适用于任意尺度，从小镇中的街区到大国中的州皆可。

现在回想一下，我们关注的是微观动机，具体而言，是个体市民的动机与行动。为研究这一点，我们假设每个街区 i都有一种可能各不相同的居住效用，而这种效用自然取决于该街区的总人口（等价地说，即人口密度）。形式上，我们有一组效用函数 f_1，f_2，……，f_n，每个函数定义在 $[0, u_i]$ 上。这里，$f_i(x_i)$ 表示当位置 i 的人口为 $x_i \in [0, u_i]$时，每位居住于该街区的个体所获得的效用。假设每个函数 f_i 均非负且为凹函数，但重要的是，它们不必是单调的。

我们指出，允许效用函数为非单调函数非常重要。尤其是，通常假设效用函数在人口密度较低时递增，而在人口密度较高时递减（例如参见 [4, 21, 33]）。这是由于人口不足和过度拥挤都会带来负外部性。例如，当总人口／人口密度过大时，过度拥挤可能造成交通拥堵、污染、噪声、犯罪等负面影响，也会因街区超额需求而推高生活成本。反过来，人口过少可能不足以支撑必要的公共服务，例如交通运输、学校、卫生服务等，也不足以支撑餐馆、商店、咖啡馆、体育设施等私营企业。

2.1 纳什均衡

当然，每个人都可以选择居住在哪个街区。这形成了一种博弈，其动态过程是人们可以依次迁往其他街区。在这种博弈中，均衡状态就是纳什均衡，即没有人能够通过迁往其他街区来提高自己的效用。

我们用一个例子来说明这一点。图1中有 $n = 2$ 个街区；我们将第一个街区称为“红色街区”，第二个称为“蓝色街区”。在这个例子中，我们假设各街区的容量相等。此外，设定 $u_1 = u_2 = 1$。再者，效用函数 f_1 和 f_2相同，并且当街区人口为$\frac{1}{2}$时达到效用峰值。最后，我们假设大都市的总人口为 $N = 1$。

在这个例子中，纳什均衡为 $(x_1, x_2) = (\frac{1}{2}, \frac{1}{2})$。这是一个可行解，因为 $N = 1 = \sum_{i=1}^{n} x_i = \frac{1}{2} + \frac{1}{2}$。此外，这也是一个纳什均衡，因为每个参与者获得的效用都是 1，达到了可能的最大值。例如，红色社区中的参与者无法通过迁往蓝色邻域来提高效用，反之亦然。事实上，对于这个具体例子，这个纳什均衡实现了可达到的最优社会福利。

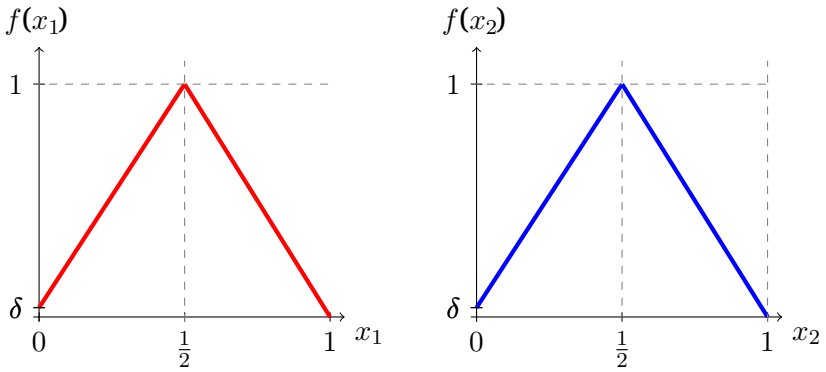


图1：一个社会福利较高的均衡实例。

2.2 社会福利与无政府状态代价

遗憾的是，这种情况并不典型。在这些城市规划博弈中，均衡／稳定状态未必能为参与者带来较高的福利。为说明这一点，我们需要以下定义。当人口分布为 $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ 时，社会福利为 $\text{wel}(\mathbf{x}) = \sum_{i=1}^{n} x_i \cdot f_i(x_i)$。可实现的最优社会福利记为 $\text{opt} = \max_{\mathbf{x} \in \mathcal{F}} \text{wel}(\mathbf{x})$，其中 $\mathcal{F}$ 表示给定城市规划实例的所有可行社区人口分布的集合。

图 2 展示了一个例子：与最优社会福利相比，均衡状态的社会福利非常低。请注意，蓝色邻域现在只比红色社区稍微更具吸引力。事实上，效用函数 f_2 现在经过三个点 $(0, 3\delta)$、$(1\,2\,, 1 + 2\delta)$ 和 $(1, 2\delta)$，而不是 $(0, \delta)$ 、 $(1\,2\,, 1)$ 和 $(1, 0)$。当 δ 较小时，这个例子与图 1 中的例子几乎相同，最优解仍为 $\mathbf{x}^* = (x_1^*, x_2^*) = (1\,2\,, 1\,2\,)$，其社会福利 $\text{opt} = \text{wel}(\mathbf{x}^*) = \sum_{i=1}^{n} x_i^* \cdot f_i(x_i^*) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (1 + 2\delta) = 1 + \delta$。然而，纳什均衡结果现在截然不同。要理解这一点，请注意，在 $(x_1^*, x_2^*) = (\frac{1}{2}, \frac{1}{2})$ 时，红色社区中的参与者效用为 1，但只要迁至蓝色邻域，就能将效用提高到（略低于） $1 + 2\delta$。此外，这种行为会自我延续，动态过程会使每位参与者都从红色社区迁往蓝色邻域。具体而言，假设当前解为 $(x_1, x_2) = (\xi, 1 - \xi)$。那么，红色社区中的参与者效用为 $f_1(\xi) = \delta + 2\xi(1 - \delta)$。迁往蓝色邻域会使其效用达到 $f_2(1 - \xi) = 2\delta + 2\xi$。但这大于 $f_1(\xi)$，因此参与者会作出这一迁移。请注意，最终的动态过程只能在唯一的纳什均衡 $\mathbf{x} = (x_1, x_2) = (0, 1)$ 处终止。但这一纳什均衡中，所有人都居住在蓝色邻域，其社会福利为 $\text{wel}(\mathbf{x}) = \sum_{i=1}^{n} x_i \cdot f_i(x_i) = 0 \cdot \delta + 1 \cdot 2\delta = 2\delta$。请注意，当 δ 较小时，这是一个糟糕的结果（尽管它是唯一的均衡）。

这个例子很好地阐释了无政府状态代价这一概念，即最优社会福利与纳什均衡的社会福利之间的最坏情况比率。[1] 如图所示，该实例的无政府状态代价为 $\frac{1+\delta}{2\delta} \approx \frac{1}{2\delta}$，当 δ 趋近于 0 时，它趋近于 ∞。事实上，

[1]无政府状态代价是博弈论和计算机科学中一个经过广泛研究的概念，用于衡量系统相对于最优性能的实际性能。该术语源于这样一种观点：在无政府状态代价较高的情况下，每个参与者都采取个人主义行动的均衡解（无政府状态）可能会导致较低的社会福利。在此类系统中，探究能否通过协调机制，或通过对底层博弈进行细微但切实可行的修改来改善结果，至关重要。

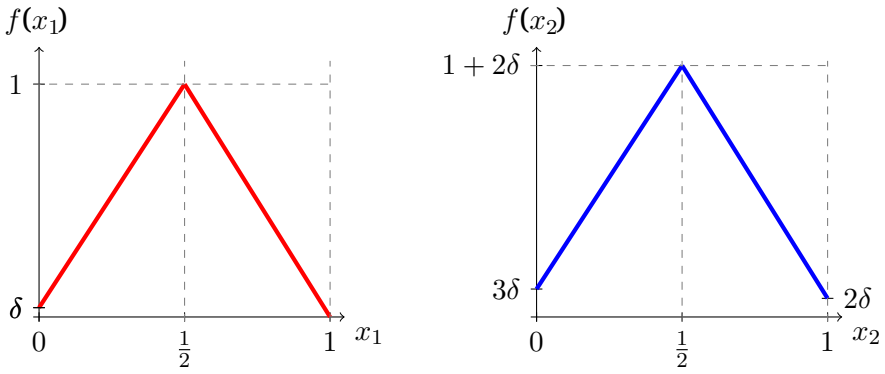


图 2：一个均衡导致社会福利低下的实例。

纳什均衡的社会福利满足 $\text{wel(x)} \leq 2\delta \cdot \text{opt}$。也就是说，当 δ较小时，纳什均衡所带来的社会福利与最优解相比微不足道。

2.3 城市规划中的布雷斯式悖论

均衡可能导致社会福利结果灾难性地恶化，这正是本研究的一大动机。具体而言，这种可能性对我们的具体研究主题——城市规划中的定向投资——有着重要影响。为说明这一点，请看图 3 中的例子。首先假设红色和蓝色街区具有相同的效用函数（红色和蓝色实线）。现在设想，经过一项小规模的定向投资后，蓝色街区略有改善，形成蓝色虚线所示的效用函数。改善之前，情况与图1相同，因此均衡时的社会福利为 1。但投资之后，情况与图 2 相同，因此均衡时的社会福利大幅下降至 2δ。由此，我们得到了布雷斯悖论的一种极端形式：街区改善所引发的微观行为，可能导致社会福利灾难性地下降。

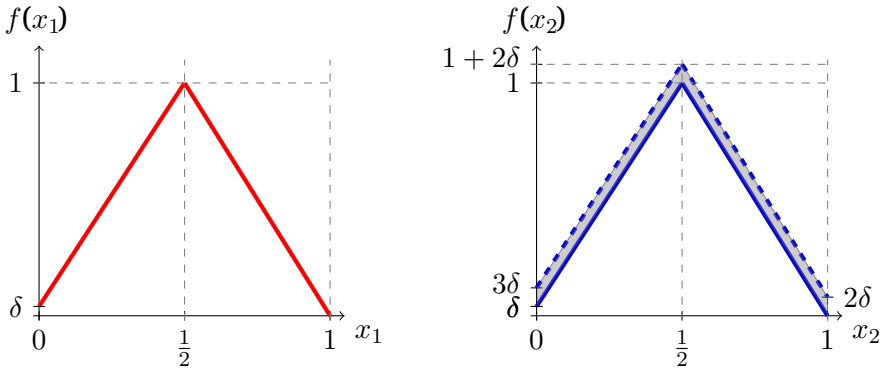


图 3：一个投资回报超负的实例。

因此，理解城市规划中微观行为造成的动态效应至关重要。最值得关注的是，这引出了以下问题：能否进行小规模、有针对性的

投资，确保由此产生的每一个均衡都具有较高的社会福利？这是我们在本文中探讨的核心问题。令人惊讶的是，我们给出了肯定的答案。

2.4 我们的结果

我们的目标是对选定的街区进行定向投资，确保由此产生的均衡下，大都市的社会福利与最优解相当。此外，我们希望这些定向投资的总成本较低。为了量化成本，我们假设对街区 i 进行定向投资，会将效用函数从 f_i 变为g_i。此外，投资成本与对称差$|f_i \oplus g_i|$的面积成正比，即两个函数之间区域的面积。因此，所有定向改进措施的总成本为 $\sum_{i=1}^{n} |f_i \oplus g_i|$。

例如，在图 3 中，定向投资的成本与蓝色邻域中阴影部分的面积成正比。请注意，该面积为 2δ ，与最优社会福利 opt $= 1 + \delta$相比微不足道。我们的主要结果是，如此小规模的定向投资始终足以确保均衡结果与最优解相当。（请注意，在图 3 中，无需投资便能实现这一点，因为原有均衡就是最优解！）

在陈述主要定理之前，我们必须解释，为什么假设投资成本与对称差的面积成正比，在城市规划场景中是合理的。首先，显然，效用函数的增幅越大（从 f_i 到 g_i），成本就越高。因此，改进成本是效用函数改进幅度（即对称差的面积）的函数。但为什么这项成本也与改进发生在 x 轴的哪个位置无关呢？原因是，改进成本大致不受街区拥堵程度的影响。具体而言，在人口稠密的街区修建学校、道路、游乐场、交通缓和措施等，成本可能比在人口稀少的街区略高，但绝不会高出一个大于常数的乘法因子。因此，我们可以假设投资成本与对称差的面积成正比。[2]

我们在第 5 节证明的主要定理指出，要确保社会福利至少达到$\epsilon \cdot$ opt，只需投入至多 $0.81\epsilon^2 \cdot$ opt。

定理1。 在任何城市规划实例中，都存在总成本至多为 $0.81\epsilon^2 \cdot$opt 的定向投资，能够确保每个由此产生的均衡，其福利都至少为 $\epsilon \cdot$opt。

举例来说，假设 $\epsilon = \frac{1}{10}$ 且我们的定向投资成本为1亿美元。新均衡下由此产生的社会福利至少为12.35亿美元。因此，如果初始均衡的福利较低，那么我们的定向投资就能带来超额正收益。事实上，值得注意的是，若通过向个人发放现金补助而非进行定向投资来获得同等福利，纳税人将付出超过10亿美元！

作为第二个示例，让我们回到贯穿全文的例子。在图4中，实线表示初始效用函数。这些函数对应于图 3 中的终端函数；我们已经看到，它们会产生一个社会福利为 2δ 的均衡，尽管最优值为 $= 1 + \delta$。假设现在我们对红色社区进行定向投资，从而将其效用函数提升至红色虚线所示的水平。这项改进的成本为 2δ，即图4中阴影部分的面积，但它

[2]此外，我们可以放宽这一假设，允许投资成本依赖于人口密度，但其差异不超过一个常数乘法因子。这样一来，本文提出的结果仍然成立，只需考虑该乘法因子。

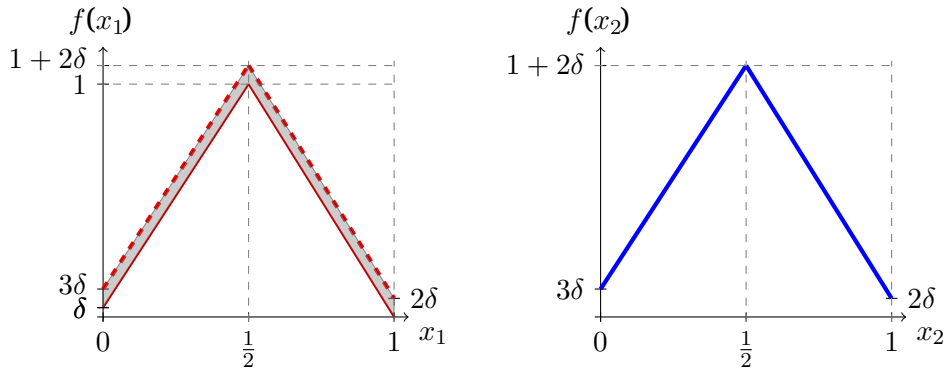


图4：一个具有超正向投资回报的实例。

立即带来唯一的纳什均衡，其社会福利为 $1+2\delta$，高于原均衡的社会福利 2δ。因此，经过一笔非常小的投资，微观行为动态就会促成一个社会福利甚至超过原先最优社会福利的结果。

再次强调，定理1的结果令人惊讶，这一点十分重要。若要让主体的总效用达到 $\epsilon\cdot$opt，另一种方法是提供补贴或减税。但这些方法的成本通常为 $\Omega(\epsilon\cdot\text{opt})$，与 $O(\epsilon^2\cdot\text{opt})$ 相比，成本大幅增加。

读者或许会问，定理1能否改进：能否以更低的成本实现这种超正向回报？答案是否定的。更具体地说，虽然常数可能略有改进，但成本函数中的主导乘法因子，也就是 ϵ^2，是紧确的，无法改进。我们将在第6节通过以下定理证明这一点。

定理2。存在一些城市规划实例，要确保每个由此产生的均衡的福利至少达到 $\epsilon\cdot\texttt{opt}$，所需定向投资的总成本至少为 $0.12\epsilon^2\cdot\texttt{opt}$。

3 背景与相关文献

我们的研究建立在博弈论、城市规划和基于智能体的建模交叉领域的丰富研究成果之上。该领域的一个核心主题是个体激励与集体福利之间的脱节——谢林在其开创性的隔离动态模型中对这一现象进行了著名探讨 [29]。谢林指出，即便个体对相似邻居只有轻微偏好，也可能导致群体结果高度隔离。

谢林最初的研究聚焦于一般图中的基于距离的邻域模型，但他也提出了有界邻域模型，在该模型中，主体评估的是整个街区，而非周边的直接环境 [29]。在这一框架下，城市被划分为容量有限的街区，主体根据各街区的整体人口或密度特征决定进入或离开。有界邻域模型揭示了策略性位置选择即使在个体偏好较弱的情况下，也会引发明显的宏观层面不连续性，例如临界点 [9, 15, 34] 以及邻域构成的突然变化。

许多后续论文以有限街区框架为基础，表明适度的利他主义 [13, 17] 和宽容度 [12]可以使系统更接近社会最优。这些结果进一步印证了一个观点：局部激励可能与整体福利严重错位，而我们希望通过定向干预来缓解这一现象。

Su等人 [31, 32] 对个体的移动方式加以限制。具体而言，他们发现，如果个体只能通过星形连接拓扑在街区之间移动，那么即使是纯粹的利己行为也能产生有效率的均衡。这些结果表明，结构性干预和行为干预都有助于缓解低效问题。

为将我们的研究置于更广泛的经济学文献背景下，我们重点介绍近期采用博弈论方法研究谢林模型改进版本的工作。Grauwin 等人 [14] 运用演化博弈论分析了该模型，表明线性效用函数能够使个体福利与集体福利保持一致，而谢林的原始模型及非对称峰值效用则往往会导致稳定但次优的隔离结果。

Grauwin 等人 [13] 对包含两个群体的谢林模型进行了小幅调整，提出了一个由同质智能体组成、且效用函数仅取决于局部密度的简化版本。在方法上，我们的研究与 Grauwin 等人 [13], 的工作最为接近，但我们将重点从涌现行为转向街区设计。我们并不假设参与者会自发改变行为，而是探讨定向投资如何促进社会期望均衡。

最近，谢林的原始模型也从博弈论视角进行了研究 [10, 11]。在该模型中，只要至少有一定比例的邻居属于同一类型，参与者就会感到满意。相关作者给出了均衡收敛性和无政府状态代价方面的结果，其中社会效用被定义为满意个体的数量。由于定义不同，这些结果无法与基于总效用的我们的结果进行比较。Kanellopoulos 等人 [18] 和 Bilò 等人 [7] 进一步研究了不同动态下的均衡低效。Agarwal等人 [2, 1] 研究了该模型的推广形式，其中包含 k 而不仅仅是两种参与者类型。

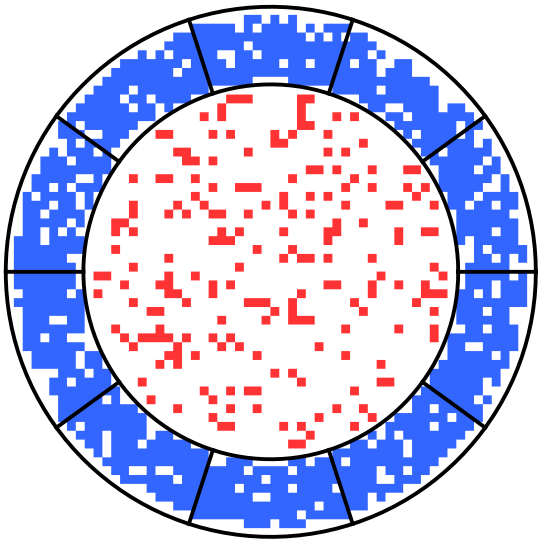
此外，我们的模型与拥塞博弈有相似之处。拥塞博弈是经济学和计算机科学中一个发展成熟的研究领域 [25, 26, 28]。在拥塞博弈的文献中，无政府状态代价已针对非原子 [28] 和原子 [3, 6, 27] 版本得到了广泛研究，也有研究探讨了旨在降低无政府状态代价的干预措施（见 [20] 及其中的参考文献）。然而，传统的拥塞博弈研究假设参与者具有单调效用结构，而我们采用的是凹效用框架。这一调整更符合受谢林启发的博弈系列的动态过程，也是我们捕捉城市邻里博弈中策略决策和资源分配所涉及的更微妙权衡所必需的。这一新视角有助于我们更好地理解并缓解智能体策略互动中的低效问题，例如利用动态过程和临界转变行为。

4 案例研究：逆转甜甜圈效应

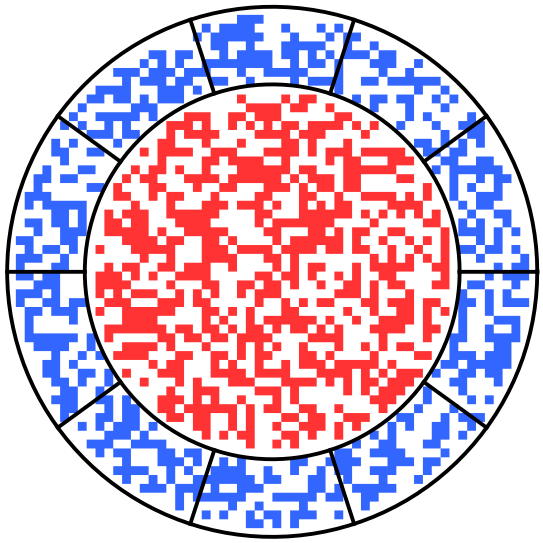
城市规划中的一个经典问题是去城市化，即人口和经济活动从人口密集的城市核心区转移到周边郊区 [5, 16, 19]。从历史上看，这一动态过程一直是在种族背景下进行研究的 [15]。事实上，在二十世纪后半叶，美国的许多大都市区都出现了这一过程，包括底特律、圣路易斯、克利夫兰、巴尔的摩和费城；这些地区的郊区人口增长，与市中心人口停滞或减少同时发生。近来，人们则从城市与郊区在服务质量、交通、住房成本、污染方面的差异，以及远程办公的普及程度等角度看待这一动态过程。导致城市核心区空心化的这一负面动态过程被称为“甜甜圈效应” [24]。

本节将说明如何利用定向投资扭转这种负面动态。图5a展示了使用我们的模型得到的甜甜圈效应空间分布示例，其中

有一个大型城市核心区和十个较小的郊区区域。在这个例子中 $n = 11$，最初城市区域与郊区的效用函数相似。尽管如此，仍形成了一种负面动态：人口从内城区迁往郊区。这个过程最终使郊区的人口密度达到 84%（以蓝色街区表示），而内城区的人口密度仅为 16%（以红色街区表示）。为便于讨论，此处无需关注得出这些统计数据的技术细节；相关效用函数和正式的数学细节请参阅附录。



(a) 起初，均衡配置中的郊区区域人口密度为 ρ_B =0。84，而内城区的人口密度较低，为 $\rho_R = 0.16$。



(b) 经过精心定向的投资后，均衡配置中郊区街区和内城区的人口密度均达到平衡，为 $\rho = 0.5$ in。

图5：一座圆形城市由一个中央内城区（红色）和一个面积相等的郊区区域（蓝色）组成，后者又细分为十个大小相同的街区。空间域离散为 60×60 网格，人口密度则以彩色填充像素表示。

然而，可以通过少量定向投资扭转这一动态。事实上，只需少量投资，略微改善内城区的效用函数，就能使人口逐步迁回城市核心区。这种正向动态会在内城区和每个郊区区域的人口密度均达到 50% 时停止。如图5b所示。我们指出，此处使用“负向”和“正向”动态这两个术语是恰当的，因为定向投资后，每位参与者的境况都得到改善（详情同样见附录）。

因此，这个例子表明，我们的模型在定性层面上捕捉到了一种空间格局，它反映了城市发展中一种稳健且有充分文献记载的现象。它还揭示了模型如何通过定向投资，帮助规避城市环境中出现的负面动态过程。

5 精准定向投资保证高福利

本节将证明我们的主要结果——定理1。

定理1. 对于任意城市规划实例，都存在总成本至多为 $0.81\epsilon^2 \cdot \mathrm{opt}$ 的定向投资，确保由此产生的每个均衡都具有至少 $\epsilon \cdot \mathrm{opt}$ 的福利。

也就是说，我们总能通过低成本的定向投资，确保任何由此产生的均衡都具有较高的社会福利。为此，有两个基本问题：如何投资？在哪里投资？“在哪里”关乎我们应该在哪些街区进行定向投资；“如何”关乎在选定的街区中应采取何种形式的定向投资。我们设计了一种算法来解决这两个问题。在第 5.1 节中，我们介绍一种有助于解决“如何”问题的方法。在第 5.2 节中，我们描述一种有效的方法，用于选出应当投资的街区。

在详细介绍算法及其数学分析之前，我们先回顾模型和问题的形式化描述。回顾一下，设有 n 个地点，其估值函数为 f_1，f_2，……，f_n，每个函数定义在 $[0, u_i]$ 上，其中 u_i 是位置 i 的容量。设有 N 个参与者，我们可以假设 $\max_{i=1}^{n} u_i \le N \le \sum_{i=1}^{n} u_i$。原因有二：第一，在实践中，$N$ 是 u_i 的事实上的上界，因为位置 i 容纳的参与者不可能多于总人口。第二，在实践中，N 是 $\sum_{i=1}^{n} u_i$ 的事实上的下界，因为所有地点加起来必须能够容纳全部人口。

此外，请记住，每个函数 f_i 都是非负且凹的，但重要的是，它不必是单调的。

我们的目标是得到满足以下条件的函数 g_1、g_2、……、g_n：

1. $\{g_1, g_2, \ldots, g_n\}$ 的均衡时社会福利至少为 $\epsilon \cdot \mathrm{opt}$。
2. 改进成本很低。具体而言，对称差满足 $\sum_{i=1}^{n} |f_i \oplus g_i| \le \epsilon^2 \cdot \mathrm{opt}$。

事实上，我们将展示如何以至多 $0.81\epsilon^2 \cdot \mathrm{opt}$ 的成本，得到一个效用至少为 $\epsilon \cdot \mathrm{opt}$ 的解。

5.1 如何做：改善地点的策略

假设我们选定一个位置进行定向投资。这些投资应采取什么形式？为了获得具有高社会福利的最终均衡，我们所采用的策略是关注极端情况，即街区的人口要么非常稀少，要么非常稠密。我们的目标是，在投资之后，无论街区人口稀少还是稠密，该街区都能为居民提供合理的（尽管可能相当有限的）效用。

现在，我们正式定义这种投资策略，并将其称为位置 i 上的 基本策略。假设 f_i 的最大值为 h_i，并且经过以下三个点：$(0, \ell_i)$、(σ_i, h_i) 和 (u_i, ℓ_i')。（中间点可能与两个端点中的某一个相同。）

改善位置 i 的初步策略如下。为位置 i 指定一个目标收益 of $\tau_i = \epsilon \cdot h_i$。如果 $\ell_i < \tau_i$，则令 α_i 满足 $f(\alpha_i) = \tau_i$，其中 $\alpha_i \le \sigma_i$；否则设为 $\alpha_i = 0$。类似地，如果 $\ell_i' < \tau_i$，则令 β_i 满足 $f(\beta_i) = \tau_i$，其中 $\beta_i \le \sigma_i$；否则设为 $\beta_i = u_i$。

给定 α_i 和 β_i，定义新函数 g_i 如下

$$g_i(x) = \begin{cases} \tau_i & \text{if } 0 \le x < \alpha_i \\ f_i(x) & \text{if } \alpha_i \le x \le \beta_i \\ \tau_i & \text{if } \beta_i < x \le u_i \end{cases}$$

我们可以通过以下引理，给出将 f_i 升级到 g_i 的总成本上界。

引理 1. 在位置 i应用初级策略，目标为 $\tau_i = \epsilon \cdot h_i$，其成本至多为 $\frac{\epsilon^2}{2} \cdot u_i \cdot h_i$。

证明。 我们来评估将 f_i 提升到 g_i的成本。根据凹性，连接 $(0, \ell_i)$ 和 (σ_i, h_i) 的线段位于 f_i 下方。该线段的斜率为 $\frac{h_i - \ell_i}{\sigma_i}$。由此可得 $\alpha_i \le (\tau_i - \ell_i) \cdot \frac{\sigma_i}{h_i - \ell_i}$。

因此，在范围 $[0, \alpha_i]$ 内，将 f_i 提升到 g_i 的成本至多为

$$\frac{1}{2} \cdot \alpha_i \cdot (\tau_i - \ell_i) \le \frac{\sigma_i}{2(h_i - \ell_i)} \cdot (\tau_i - \ell_i)^2. \tag{1}$$

类似地，根据凹性，连接 (σ_i, h_i) 和 (u_i, ℓ_i') 的线段位于 f_i 下方。该线段的斜率为 $-\frac{h_i - \ell_i'}{u_i - \sigma_i}$。由此可得 $u_i - \beta_i \le (\tau_i - \ell_i') \cdot \frac{u_i - \sigma_i}{h_i - \ell_i'}$。

将 f_i 改进到 g_i （范围为 $[\beta_i, u_i]$ ）的成本至多为

$$\frac{1}{2} \cdot (u_i - \beta_i) \cdot (\tau_i - \ell_i') \le \frac{(u_i - \sigma_i)}{2(h_i - \ell_i')} \cdot (\tau_i - \ell_i')^2. \tag{2}$$

因此，根据（1） 和（2），对地点 i 进行这两项改进的总成本至多为

$$\begin{aligned} \frac{\sigma_i}{2(h_i - \ell_i)} \cdot (\tau_i - \ell_i)^2 + \frac{(u_i - \sigma_i)}{2(h_i - \ell_i')} \cdot (\tau_i - \ell_i')^2 &\le \frac{\sigma_i}{2h_i} \cdot \tau_i^2 + \frac{(u_i - \sigma_i)}{2h_i} \cdot \tau_i^2 \\ &= \frac{u_i \cdot \tau_i^2}{2h_i} \\ &= \frac{\epsilon^2}{2} \cdot u_i \cdot h_i. \end{aligned} \tag{3}$$

这里的不等式成立是因为 $\tau_i \le h_i$，第二个等式成立是因为 $\tau_i = \epsilon \cdot h_i$。 □

5.2 位置：选择要改进的位置

现在，我们来探讨一个复杂的问题：应当在哪里进行定向投资。问题在于，试图改进每一个位置的成本会过于高昂。因此，我们必须谨慎选择哪些地点采用初步策略进行改进。

为选择地点，我们的算法将问题分为两种情况。如果某个位置 ℓ^*满足 $u_{\ell^*} \cdot h_{\ell^*} \ge \phi \cdot \mathsf{opt}$，则称其为关键位置，其中 $\phi = \frac{\sqrt{5}-1}{2} \approx 0.618$ 为黄金比例共轭数。城市规划实例中要么存在关键位置，要么不存在任何关键位置。

如果出现关键位置的情况（尽管可能性不大），那么选择问题就很简单。我们必须将投资集中在关键位置。让我们将这一方法形式化，并分析其成本和性能。

(1) 存在一个关键位置 ℓ^*。

令 $u_{\ell^*} \cdot h_{\ell^*} = \lambda \cdot \mathsf{opt}$，其中 $\lambda \geq \phi$。在这种情况下，我们的策略很简单：只改进位置 ℓ^*。具体而言，我们在位置 ℓ^* 采用初步策略，使用目标收益 $\tau_{\ell^*} = \epsilon^* \cdot h_{\ell^*}$，其中 $\epsilon^* = \frac{1}{\lambda} \cdot \epsilon$。

引理 2。 如果存在关键位置，则总改进成本至多为 $0.81 \cdot \epsilon^2 \cdot \mathtt{opt}$，且任何所得均衡的社会福利至少为 $\epsilon \cdot \mathtt{opt}$。

证明。 考虑在位置 ℓ^* 采用初步策略并使用目标收益 $\tau_{\ell^*} = \epsilon^* \cdot h_{\ell^*}$后得到的任意纳什均衡 $\mathbf{x}$。如果 $x_{\ell^*} = u_{\ell^*}$，则所得社会福利至少为

$$u_{\ell^*} \cdot \tau_{\ell^*} \;=\; \epsilon^* \cdot u_{\ell^*} \cdot h_{\ell^*} \;=\; \frac{1}{\lambda} \cdot \epsilon \cdot u_{\ell^*} \cdot h_{\ell^*} \;=\; \epsilon \cdot \mathtt{opt}. \tag{4}$$

另一方面，假设 $x_{\ell^*} < u_{\ell^*}$。那么根据纳什均衡条件，每个参与者的效用至少为 τ_{ℓ^*}，否则它就会通过移动到位置 ℓ^*来提高自己的效用。因此，社会福利至少为

$$N \cdot \tau_{\ell^*} \;=\; N \cdot \epsilon^* \cdot h_{\ell^*} \;=\; N \cdot \frac{1}{\lambda} \cdot \epsilon \cdot h_{\ell^*} \;\geq\; \frac{1}{\lambda} \cdot \epsilon \cdot u_{\ell^*} \cdot h_{\ell^*} \;=\; \epsilon \cdot \mathtt{opt}. \tag{5}$$

上文中的不等式成立，因为 $N \geq \max_{i=1}^{n} u_i \geq u_{\ell^*}$。

由 (4) 和 (5) 可知，纳什均衡的社会福利至少为 $\epsilon \cdot \mathsf{opt}$。要获得这一福利保证，所需的成本是多少？根据引理 1，所需成本至多为

$$\begin{aligned}
\frac{(\epsilon^*)^2}{2} \cdot u_{\ell^*} \cdot h_{\ell^*} &= \frac{\epsilon^2}{2\lambda^2} \cdot u_{\ell^*} \cdot h_{\ell^*} \\
&= \frac{\epsilon^2}{2\lambda} \cdot \mathtt{opt} \\
&\leq \frac{\epsilon^2}{2\phi} \cdot \mathtt{opt} \\
&= \frac{1+\phi}{2} \cdot \epsilon^2 \cdot \mathtt{opt}.
\end{aligned} \tag{6}$$

这里，最后一个不等式成立，因为 ϕ 是黄金比例共轭数，因此 $\phi(1+\phi) = 1$。

因此，只需付出至多 $\frac{1+\phi}{2} \cdot \epsilon^2 \cdot \mathsf{opt} \approx 0.81 \cdot \epsilon^2 \cdot \mathsf{opt}$ 的成本，我们就能得到效用至少为 $\epsilon \cdot \mathsf{opt}$ 的解。 □

在没有关键位置的情况下，选择要改进哪些地点要微妙得多。

(2) 不存在关键位置。

由此可知，每个位置 i 都具有如下性质：$u_i \cdot h_i < \phi \cdot \mathsf{opt}$。对地点排序，使得 $h_1 \geq h_2 \geq \cdots \geq h_n$。令

$$k^* = \min_k \{k : \sum_{i=1}^{k} u_i \cdot h_i \geq \mathtt{opt}\}.$$

注意到 k^* 作为最优解 $\leq \sum_{i=1}^{n} u_i \cdot h_i$ 存在。令 λ 满足

$$\sum_{i=1}^{k^*} u_i \cdot h_i \;=\; \lambda \cdot \mathtt{opt}. \tag{7}$$

注意， $1 \leq \lambda$ 和

$$\lambda < 1 + \phi, \tag{8}$$

这由 $\sum_{i=1}^{k^*-1} u_i \cdot h_i <$ 最优得出，并根据情形假设， $u_{k^*} \cdot h_{k^*} < \phi \cdot \mathsf{op}$ t.

现在我们的策略是只改进位置 1 到 k^*。不改进位置 $k^* + 1$ 到 n。具体来说，对于每个 $1 \leq i \leq k^*$，我们在位置 i应用以$\tau_i = \epsilon \cdot h_i$ 为目标的初步策略。

引理 3. 如果不存在关键位置，那么总改进成本至多为 $0.81 \cdot \epsilon^2 \cdot \mathtt{opt}$，且任何由此产生的均衡的社会福利至少为 $\epsilon \cdot \mathtt{opt}$。

证明。 令 $\mathbf{x}$ 为一个由此产生的纳什均衡。令 $j^* \leq k^*$ 为满足 $x_{j^*} < u_{j^*}$的第一个位置（如果存在这样的位置）。

如果不存在这样的位置，则令 $j^* = k^* + 1$。由于位置 1 到 k^* 均已满，纳什均衡的总效用至少为

$$\begin{aligned} \sum_{i=1}^{k^*} u_i \cdot \tau_i &= \epsilon \cdot \sum_{i=1}^{k^*} u_i \cdot h_i \\ &\geq \epsilon \cdot \mathtt{opt}. \end{aligned} \tag{9}$$

另一方面，假设 $j^* \leq k^*$。那么在此纳什均衡中，位于位置 $i > j^*$ 的每个参与者的效用至少为 τ_{j^*}；否则它就会移至位置 j^*。

同样，由于位置 1 到 $j^* - 1$ 均已满，纳什均衡的总效用至少为

$$\begin{aligned} \sum_{i=1}^{j^*-1} u_i \cdot \tau_i + (N - \sum_{i=1}^{j^*-1} u_i) \cdot \tau_{j^*} &= \epsilon \cdot \left(\sum_{i=1}^{j^*-1} u_i \cdot h_i + (N - \sum_{i=1}^{j^*-1} u_i) \cdot h_{j^*} \right) \\ &\geq \epsilon \cdot \left(\sum_{i=1}^{k^*-1} u_i \cdot h_i + (N - \sum_{i=1}^{k^*-1} u_i) \cdot h_{k^*} \right) \\ &\geq \epsilon \cdot \mathtt{opt}. \end{aligned} \tag{10}$$

上面的第一个不等式成立，是因为 h_i 按降序排列。此外，由于 h_i是递减的，因此必有 $\mathsf{opt} \leq \sum_{i=1}^{k^*-1} u_i \cdot h_i + (N - \sum_{i=1}^{k^*-1} u_i) \cdot h_{k^*}$，从而得到第二个不等式。

由 (9) 和 (10) 可知，纳什均衡的社会福利至少为 $\epsilon \cdot \mathsf{opt}$。获得这一福利保证所付出的成本是多少？回想一下，我们只改进位置 1 到 k^* （即对 $1 \leq i \leq k^*$使用 f_i，而非 g_i ）。对于位置 $k^* + 1$ 到 n ，我们不作任何改进（即对 $k^* + 1 \leq i \leq n$继续使用 f_i），因此这些位置不会产生任何成本。

因此，根据式（3），我们对地点 1 至 k^* 所做改进的成本至多为

$$\begin{aligned} \sum_{i=1}^{k^*} \frac{\epsilon^2}{2} \cdot u_i \cdot h_i &= \frac{\epsilon^2}{2} \cdot \sum_{i=1}^{k^*} u_i \cdot h_i \\ &= \frac{\epsilon^2}{2} \cdot \lambda \cdot \mathtt{opt} \\ &\leq \frac{1 + \phi}{2} \cdot \epsilon^2 \cdot \mathtt{opt}. \end{aligned} \tag{11}$$

这里，等号由式（7）得出，不等式则由式（8）得出。

因此，我们再次得到一个效用至少为 $\epsilon\cdot\text{opt}$ 的解，而其成本至多仅为 $\frac{1+\phi}{2}\cdot\epsilon^2\cdot\text{opt}\approx 0.81\epsilon^2\cdot\text{opt}$。□

这引出了我们的主要结果。

定理 T 的证明 定理 *1*。该定理是引理 2 和引理的直接推论 3. □

6 一个最优保证

读者可能会问，定理1所提供的保证是否可以改进。答案是否定的：该定理给出了可达到的最强保证。具体而言，如果我们希望确保每个最终均衡的社会福利至少为 $\epsilon\cdot\text{opt}$，那么就必须进行总成本为 $\Omega(\epsilon^2\cdot\text{opt})$ 的投资。因此，ϵ^2 项是必要的，定理1在常数因子范围内是紧的。以下定理对此作出了说明。

定理2。存在这样的城市规划实例：若要确保每个最终形成的均衡的福利至少为 $\epsilon\cdot\text{opt}$，则所需定向投资的总成本至少为 $0.12\epsilon^2\cdot\text{opt}$。

证明。要证明该定理，只需给出一个具有此性质的例子。为此，我们构造一个包含 $n=2$ 个地点的实例，每个地点的容量为 $u_1=u_2=1$，参与者数量为 $N=1$。令 f_1 为经过点 (0，1) 和 (1，0) 的分段线性函数。因此 $h_1=1$。令 f_2 为经过点 (0，0) 和 (1，0) 的常函数。因此 $h_2=0$。

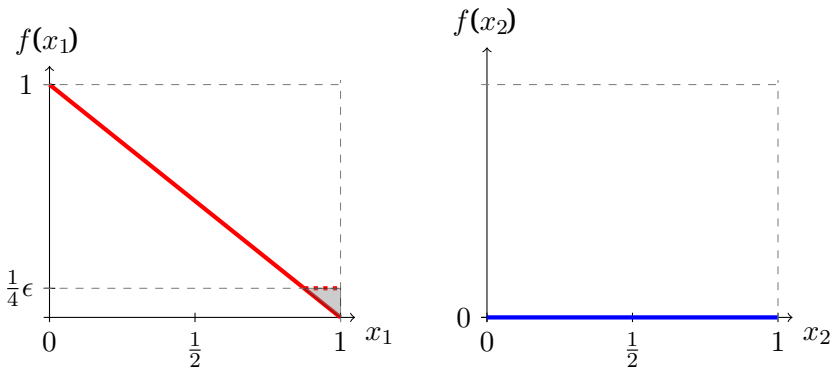


图6：下界示例

令 $x_1=\lambda$ 和 $x_2=1-\lambda$，可得社会福利为 $\lambda(1-\lambda)$，当 $\lambda=\frac{1}{2}$ 时该福利达到最大。因此 $\text{opt}=\frac{1}{4}$。

现在假设我们想要构造一个社会福利为 $\epsilon\cdot\text{opt}=\frac{1}{4}\epsilon$ 的纳什均衡。为此，读者可以注意到，改善位置 2 比改善位置 1 成本更高。因此，成本最低的解是只改善位置 1。为此，我们设定：

$$g_1(x)=\begin{cases} f_i(x) & \text{if } 0\le x\le 1-\frac{1}{4}\epsilon \\ \frac{1}{4}\epsilon & \text{if } 1-\frac{1}{4}\epsilon < x\le 1 \end{cases}$$

注意，现在存在一个唯一的纳什均衡：每个参与者都移动到位置 1，并获得 $\frac{1}{4}\epsilon$ 的收益。由于有 $N = 1$ units 的参与者，因此社会福利为 $\epsilon \cdot$opt，正如所期望的那样。

改善 f_1 至 g_1 的总成本为

$$\frac{1}{2}\left(\frac{1}{4}\epsilon\right)^2 = \frac{1}{32}\epsilon^2 = \frac{1}{8}\epsilon^2 \cdot \texttt{opt}.$$

因此，最优解的最低成本为 $\Omega(\epsilon^2 \cdot$opt$)$。 □

请注意，定理 1 和定理 2 中的乘法常数分别为 0.81 和 $1\ 8$。缩小这些上界与下界之间的差距是一个有趣的开放性问题。更乐观地说，是否有可能完全弥合这一差距，从而为我们的超正结果得到确切的乘法常数?

7 结论

利用谢林的广义有界邻域模型 [29], ，我们证明了：当前的社区规划要么已经能够提供与可实现的最优社会福利相当的社会福利，要么存在能够产生超正回报的目标性社区投资。这一结果构成了理论上的“概念验证”。这立即引出了两个关键问题。第一，这一结果能否扩展到更加现实的城市规划模型? 第二，这一结果能否在实验环境中得到验证?

参考文献

[1] A. Agarwal、E. Elkind、J. Gan、A. Igarashi、W. Suksompong 和 A. Voudouris。图上的谢林博弈。人工智能，301:103576，2021年。[2] A. Agarwal、E. Elkind、J. Gan 和 A. Voudouris。图上谢林博弈中的交换稳定性。载于第*34*届人工智能会议论文集（*AAAI*），第1758–1765页，2020年。[3] S Aland、D. Dumrauf、M. Gairing、B. Monien 和 F. Schoppmann。多项式拥塞博弈的精确无政府状态代价。*SIAM*计算杂志，40(5):1211–1233，2011年。[4] J. Barr 和 T. Tassier。隔离与策略性邻里互动。《东部经济学杂志》，34:408–503，2008年。[5] N. Baum-Snow。高速公路是否导致了郊区化? 《经济学季刊》，122(2):775–805，2007年。[6] K. Bhawalkar、M. Gairing 和 T. Roughgarden。加权拥塞博弈：无政府状态代价、普适最坏情形示例及紧确性。*ACM*经济学与计算汇刊，2(4):14:1–14:23，2014年。[7] D. Bilò、V. Bilò、P. Lenzner 和 L. Molitor。交换型谢林博弈中的拓扑影响与局部性。《自主智能体与多智能体系统》，36(2):47，2022年。

[8] D. Braess。交通规划中的一个悖论。《企业研究与运筹学》，12(1):258–268，1968年。[9] D. Card、A. Mas 和 J. Rothstein。临界转变与隔离动态。《经济学季刊》，123:177–218，2008年。[10] A. Chauhan、P. Lenzner 和 L. Molitor。具有策略性智能体的谢林隔离。载于第*11*届国际算法博弈论研讨会（*SAGT*）论文集，第137–149页，2018年。[11] H. Echzell、T. Friedrich、P. Lenzner、L. Molitor、M. Pappik、F. Schöne、F. Sommer 和 D. Stangl。策略性谢林隔离的收敛性与困难性。载于第*15*届国际网络与互联网经济学会议（*WINE*）论文集，第156–170页，2019年。[12] L. Gauvin、J. Vannimenus 和 J-P. Nadal。谢林隔离模型的相图。《欧洲物理学杂志*B*》，70:293–304，2009年。[13] S. Grauwin、E. Bertin、R. Lemoy 和 P. Jensen。集体动态与个体动态之间的竞争。《美国国家科学院院刊》，106(49):20622–20626，2009年。[14] S. Grauwin、F. Goffette-Nagot 和 P. Jensen。居住隔离的动态模型：解析解。《公共经济学杂志》，96(1-2):124–141，2012年。[15] M. Grodzins。大都市隔离。芝加哥大学出版社，1957年。[16] V. Henderson 和 J-F. Thisse。《区域与城市经济学手册：城市与地理》，第4卷。爱思唯尔，2004年。[17] P. Jensen、T. Matreux、J. Cambe、H. Larralde 和 E. Bertin。利他主义者在谢林隔离模型中的巨大催化效应。《物理评论快报》，120(20):208301，2018年。[18] P. Kanellopoulos、M. Kyropoulou 和 A. Voudouris。修正谢林博弈。《理论计算机科学》，880:1–19，2021年。[19] P. Mieszkowski 和 E. Mills。大都市郊区化的成因。《经济展望杂志》，7(3):135–147，1993年。[20] D. Paccagnan 和 M. Gairing。在拥塞博弈中，税收可实现最优近似。《运筹学》，72(3):966–982，2024年。[21] R. Pancs 和 N. Vriend。重新审视谢林的空间邻近隔离模型。《公共经济学杂志》，91:1–24，2007年。[22] B. Plaza。毕尔巴鄂古根海姆博物馆的投资回报。《国际城市与区域研究杂志》，30(2):452–467，2006年。[23] B. Plaza。毕尔巴鄂效应（毕尔巴鄂古根海姆博物馆）。《博物馆新闻》，86(5)，2007年。[24] A. Ramani 和 N. Bloom。甜甜圈效应：COVID-19如何影响房地产。*SIEPR*政策简报，一月，第1–8页，2021年。

[25] R. Rosenthal. 一类具有纯策略纳什均衡的博弈。《国际博弈论杂志》，2:65–67，1973年。[26] T. Roughgarden.《自私路由与无政府状态代价》。麻省理工学院出版社，2005年。[27] T Roughgarden. 无政府状态代价的内在稳健性。《美国计算机学会会刊》，62(5):32:1– 32:42，2015年。[28] T. Roughgarden 和 É. Tardos. 自私路由有多糟糕？《美国计算机学会会刊》，49(2):236–259，2002年。[29] T. Schelling. 隔离动态模型。《数学社会学杂志》，1(2):143–186，1971年。[30] R. Steinberg 和 W. Zangwill. 布雷斯悖论的普遍性。《交通科学》，17(3):301–318，1983年。[31] G. Su、Q. Xiong 和 Y. Zhang. 带区块的谢林模型中底层星形拓扑的引人注目效应。《物理评论*E*》，102(1):012317，2020年。[32] G. Su 和 Y. Zhang. 星形底层拓扑对谢林元种群模型中隔离现象的显著抑制。《欧洲物理学杂志*B*》，96(7):91，2023年。[33] J. Zhang. 一个人人都主张融合的世界中的居住隔离。《经济行为与组织杂志》，54:533–555，2004年。[34] J. Zhang. 临界转变与居住隔离：统一谢林模型。《区域科学杂志》，51:167–193，2011年。

附录

在本附录中，我们给出第4节所述甜甜圈效应示例的形式化细节。该示例由1个容量为1000的红色大型内城区邻里，以及10个容量各为100的蓝色小型郊区邻里组成。（为便于理解，读者可以假设这些单位以千为计量单位，因此内城区的最大人口容量为一百万，而十个郊区的最大总人口容量合计为一百万。）红色邻里和蓝色邻里的效用函数如下：

$$f_R(x) = \begin{cases} 25 + \frac{3}{20}x & \text{if } 0 \le x \le 500 \\ 200 - \frac{1}{5}x & \text{if } 500 \le x \le 1000 \end{cases} \qquad f_B(x) = \begin{cases} 26 + \frac{3}{20}x & \text{if } 0 \le x \le 50 \\ 101 - \frac{43}{31}(x-50) & \text{if } 50 \le x \le 81 \\ 301 - 3x & \text{if } 81 \le x \le 100 \end{cases}$$

这些效用函数如图7所示。请注意，除了蓝色邻里具有折点向下斜率之外，各街区都相似。注意，折点位于（81，58）处。

假设这座城市的总人口为 1000 个单位（即一百万人），并假设内城区和郊区街区的初始容量均为 50%。这个例子令人担忧，因为随着时间推移，内城区的参与者会不断迁往郊区，直到达到均衡。在这个唯一的均衡状态下，内城区人口稀少

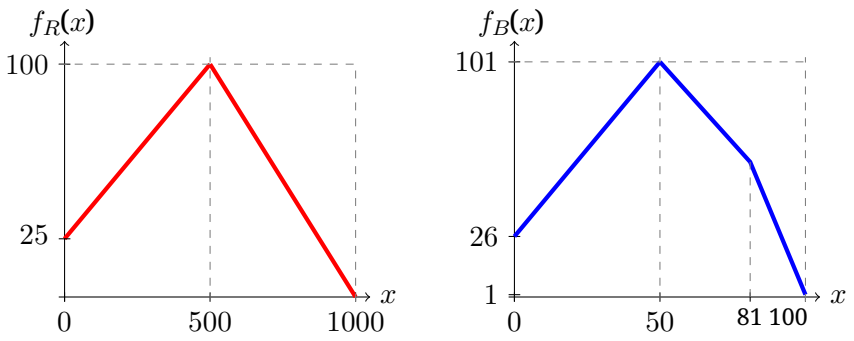


图7：一个存在城市人口外迁的实例。

（容量为 16%），而郊区人口稠密（容量为 84%）。尽管郊区和内城区的质量非常相似——它们的效用函数也表明了这一点——但仍然如此。

为证明这一点，假设红色社区中有 $10x$ 个单位的参与者，而每个蓝色邻域中有 $100-x$ 个单位的参与者；我们指出，在任何均衡状态下，每个蓝色邻域中的数量都将相同。现在，对于 $50 \geq x \geq 19$，我们有

$$\begin{aligned} f_B(100-x) - f_R(10x) &= \left(101 - \frac{43}{31} \cdot ((100-x) - 50)\right) - \left(25 + \frac{3}{20} \cdot 10x\right) \\ &= 76 + \frac{43}{31} \cdot (50 - x) - \frac{3}{2}x \\ &= 145.355 - 2.887x \\ &\geq 145.355 - 2.887 \cdot 50 \\ &> 0 \end{aligned}$$

因此，只要 $x \geq 19$，参与者就能通过从内城区的红色社区迁移到郊区蓝色邻域来获得更高的效用。事实上，这一过程会持续到 $x = 16$。具体而言，对于 $19 \geq x > 16$，我们有

$$\begin{aligned} f_B(100-x) - f_R(10x) &= (301 - 3 \cdot (100-x)) - \left(25 + \frac{3}{20} \cdot 10x\right) \\ &= \frac{3}{2}x - 24 \\ &> \frac{3}{2} \cdot 16 - 24 \\ &= 0 \end{aligned}$$

注意到，该过程在此终止 $x = 16$，此时我们得到唯一的纳什均衡：内城区拥有 $x_R = 10 \cdot 16 = 160$ 人口单位，每个郊区邻里拥有 $x_B = 100 - 16 = 84$ 人口单位。

这个均衡确实产生了图5a所示的甜甜圈图景。如何通过少量定向投资来纠正这种甜甜圈效应？为此，考虑稍微改善红色的内城街区后会出现什么情况：

$$g_R(x) = \begin{cases} 32 + \frac{69}{500}x & \text{if } 0 \leq x \leq 500 \\ 202 - \frac{101}{500}x & \text{if } 500 \leq x \leq 1000 \end{cases}$$

这条新的效用函数以图8中的虚线表示。灰色阴影区域表示这项改进所对应的投资成本。

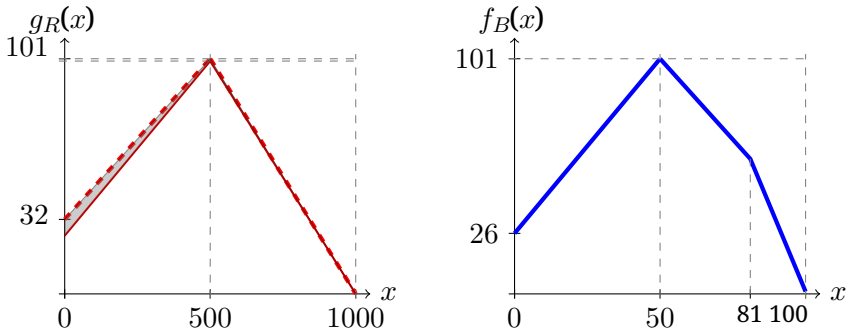


图8：改善内城区。

我们认为，这一改善的作用是产生一种反向动态，使参与者从郊区迁回内城区。当内城区邻里和每个郊区邻里都恰好达到 50% 的容量时，这一动态便会终止。为说明这一点，首先考虑 $x \leq 19$ 的情况。此时我们有

$$\begin{aligned} g_R(10x) - f_B(100-x) &= \left(32 + \frac{69}{500} \cdot 10x\right) - (301 - 3 \cdot (100 - x)) \\ &= 31 - \frac{81}{50}x \\ &\geq 31 - \frac{81}{50} \cdot 19 \\ &> 0 \end{aligned}$$

由此可见，当 $x \leq 19$时，人们会从郊区迁回内城区。当 $x \geq 19$ 时会发生什么？那么，对于 $19 < x < 50$，我们有

$$\begin{aligned} g_R(10x) - f_B(100-x) &= \left(32 + \frac{69}{500} \cdot 10x\right) - \left(101 - \frac{43}{31} \cdot ((100-x) - 50)\right) \\ &\geq 0.354887 - 0.007097x \\ &\geq 0.354887 - 0.007097 \cdot 50 \\ &> 0 \end{aligned}$$

这意味着，主体从郊区迁回内城区仍能获得更高的效用。这个过程会持续下去，直到 $x = 50$ ，最终达到唯一的纳什均衡：每个街区的容量都恰好为50%。因此，定向投资引发了连锁反应，直到居住在内城区的人数与居住在所有郊区的人数总和相等。如前所述，图5b所示的最终均衡扭转了甜甜圈效应。