

自饮性：制度化认知的成功何以系统性地耗损其认知能力

金华 著·依王德生《SDE本体论》·发表于2026年7月28日

摘要

当代社会的预测失灵——从金融模型到民意调查——常被归咎于外部混沌、认知局限或技术瑕疵。本文拒绝这些解释共享的预设（“预测是对外部世界的智识穿透”），并以此为起点推进一个硬转向：预测不是一个认知问题，而是一个制度性的物质架构问题。在知识社会学与科学技术研究传统已论证了“制度制造可见性”的基石上，本文将论证推入一个新的辨别面：制度化认知在其成功运行中，不仅制造了可见性，更会结构性地喂养那些系统性脱逃其计量语法的不可读地带。制度化的可读性，与被它喂养壮大的不可读性，不是此消彼长的外部对立，而是同一枚硬币的两面。本文将这一机制命名为自饮性（autopotivity）——制度化认知在成功中将自身赖以成立的物质语法、公共信息管道与参照系转化为耗损其认知能力的反向系统的养料。与古德哈特定律、免疫自噬理论、卢曼式自创生以及操演性文献的系统性边界划定表明，自饮性并非对这些已有概念的重述，而是切开了它们均无法整体装下的一个辨别面：反向系统并不攻击母体，它只是以母体为食，在持续汲取中使母体的认知能力逐渐衰竭。本文最后给出可裁决的经验判据，并以一组变异案例展示机制的条件性而非决定论，正面回应效率论与能动性理论的反驳。

关键词

预测失灵；自饮性；物质语法；制度基础设施；可裁决性；古德哈特定律；操演性

一、引言：被误置的问题

过去二十年间，一种特定的认知耗损在全球多个关键制度领域持续蔓延：金融风险模型在系统性危机面前屡屡失效；选举民调在投票箱前反复出现重大偏差；科技预测对于颠覆性发明的路径一再失准；情报机构对重大地缘政治变局的预警频频落空。面对这一系列可观察的耗损，主流公共话语和学术解释迅速分化为三种彼此竞争但共享同一地基的叙述：外部混沌论（Taleb, 2007 的“黑天鹅”），认知缺陷论（行为经济学所揭示的启发式偏误与过度自信），以及技术待补论（持续呼吁更大规模的模型、更精密的算法和更丰富的数据）。

这三种叙述虽在具体归因上争吵不休，却从未质疑过一个共同的预设：预测在本质上是一种智识对外部混沌的穿透。在这一预设之上，“失灵”被自然地理解为穿透力的不足——要么是外部混沌过于凶猛，要么是内部认知过于脆弱，要么是手中的工具尚不够锋利。问题的提问方式本身，已经提前锁定了答案的形态：更多的理性、更好的模型、更全的数据。

本文试图撬开这个地基的一角。知识社会学与科学技术研究的经典传统早已给出了另一套观察世界的方式。阿兰·德罗西埃（Desrosières, 1993）揭示了现代统计数字并非对“现实”的中性发现，而

是通过分类、编码与制度管道制造出可供治理的可见性。西奥多·波特（Porter, 1995）论证了量化客观性并非方法论的自动产物，而是一项昂贵的、制度化的社会信任工程。詹姆斯·斯科特（Scott, 1998）剖析了现代国家赖以运转的“可读性”简化机制——标准化的度量、统一的地籍、永久的姓氏——如何同时构成了国家认知能力的基础与盲区。保罗·爱德华兹（Edwards, 2010）在《庞大机器》中直接展示：气象预测的可靠性并非源于科学模型的优越性，而是源于一套跨国制度网络对数据生产、传输和标准化的长期维护。这些工作联手将一颗钉子钉牢：预测不是智识穿透混沌，而是制度在混沌之上搭建脚手架。

但是，这套脚手架搭成之后的事情，这些经典论述普遍没有追问到底。它们出色地论证了制度如何“制造”现实，却较少追问：当这套制造过程在其自身的成功运行中，开始系统性地喂养那些瓦解它的力量时，会发生什么？更精确地说，预测能力的耗损不仅仅是“制度不在场”的证据——它更可能是在场制度太成功、太深入人心、太被行动者内化进自身的策略计算之后，所必然结算出的一种反向生成物。

在展开全文论证之前，有必要澄清本文使用的“制度”概念的具体理论位置。在STS传统与历史制度主义之间，存在一条隐而未发的交汇带：前者提供了分析“制度制造可见性”的物质工具，后者揭示了“制度塑造激励结构”的行为机制。本文的制度概念正是在这一交汇带上展开——既从STS传统汲取对物质性、计量基础设施与可见性政治的敏感，也从制度主义传统汲取对激励结构、机会成本与策略性反应的关注。二者在此处衔接的必要性在于：只有当“制度制造可见性”这一物质过程与“被看见会改变行为的激励结构”相遇时，本文所要追踪的那条反向生成链才会被全面地激活。

本文的论证将沿着以下路径推进。第二章以气象学为锚，正面展示一套制度化的认知体系如何在成功中维持稳定，并恰恰在对比中揭示这种稳定所依赖的不可复制的条件，由此逼出“成功何以可能不喂养失明”这一核心问题。第三章解剖金融监管史上“计量-套利”的耦合演化，论证规避行为不是外部的败德或疏漏，而是被可读性的成功从其自身内部逼出的结构对应物。第四章追踪公共信息管道如何被自己所浇灌出的私有算法从内部取代——并非被外部冲击击垮，而是被自己成功所培育的更高效替代者掏空。第五章处理预测模型的自反性陷阱：一个被广泛采纳的模型，会在其成功的操演中，从“表演现实”滑向“吞噬现实”，切断其与外部参照系之间的神经。

这四章的论证将共同指向一个无法再被旧有概念所整体容纳的东西。第六章将撤销那个从未被质疑的底层先验——“制度化的秩序本身是中性工具”，并从一个更根本的三元结构（秩序的生产、流通与反馈）出发，将三条分散的机制链锚定在同一颗逻辑硬核之上，令“自饮性”从一个敏锐的观察凝固为一个具有结构解释力的中层概念。第七章将此概念置于其最强大的四个理论对手的探照灯下——古德哈特定律、埃斯波西托的免疫自噬论、卢曼的系统自创生以及操演性文献——通过代理坍缩分析，证明自饮性并非对任一对手的修辞换皮，而是切开了它们各自只握有一侧的那条辨别之缝，并给出可供裁决的经验判据。第八章正面回应来自效率学派与能动性理论的最强反驳，并引入一组变异案例以展示机制的条件性，从而避免滑向决定论。结论章拒绝任何“回到夯土”的怀旧呼吁——那在自饮性的逻辑面前是一个概念上的空动作——转而追问：当旧的制度化认知体系在结构性自饮中持续沉降时，那些在不同认知组织方式之中的微弱尝试，究竟是被危机逼出的新路径，还是自饮性在更高阶维度上的猎物？

二、成功的骨架：气象学如何维持可读性而不自噬

若要正面理解“制度化认知的成功是如何可能的”，没有比现代气象学更干净的观察样本。

今天，一个普通人打开手机就能获取未来一周的天气预报，其准确率之高已令我们习以为常。然而，这项能力并非从天而降，而是建立在耗费数代人、无数国家公共预算、以及在冷战中跨越意识形态铁幕才得以锻造的一套“物质语法”脊骨之上。所谓物质语法，指的是一套使混沌现实得以被持续计量、被横向比较、被整合进统一预测模型的基础设施系统：它包括标准化的度量衡（温度、气压、风速的统一定义与仪器校准规范）、不间断的全球数据采集节点（地面观测站、高空气球、卫星），以及能将离散信息凝聚为整体认知图景的制度管道（数据交换协议、传输标准、共享义务）。之所以称之为“语法”而非更常见的“基础设施”，正是为了强调其规范性特征：一套物质语法一旦建立，就反过来定义了何谓“正确的知识”——不符合其标准的信息在专业意义上是不算数的。正如一套语言的语法既使言说成为可能也同时划定了什么可被言说，预测的物质语法既使计量成为可能也划定了什么可被计量。

这根脊骨的构造史并非单纯的技术发明史，而是一场漫长的制度博弈史。十九世纪中叶，航海业与农业对风暴预警的迫切需求撞上了一堵硬墙：天气是跨越大陆与海洋的连续现象，任何单一国家的观测网络都只能捕捉其极小的一片切面。没有统一的观测时刻、没有标准化的仪器校准、没有跨国的无偿数据共享机制——“天气预报”在物理上是不可能实现的。这堵墙不是被某个天才的气象学家独自推倒的，而是在数十年国际谈判、条约签署与国会预算审批的接力中被一块块凿开的。1873年国际气象组织成立；二战后，1947年世界气象公约在华盛顿签署；1951年世界气象组织（WMO）正式成为联合国专门机构。成员国——无论其意识形态阵营——承担了条约义务：维持地表和高空观测网络，按统一标准持续采集数据，并无偿向全球交换。在柏林封锁与古巴导弹危机的最紧张时刻，跨越铁幕的气象数据交换从未中断过一天。这是一项耗费惊人的公共工程：截至2020年代，全球超过一万个地面观测站、数千个高空气象站、数十颗卫星、数百台超级计算机在支撑这一网络，年度维护费用以百亿美元计，由各国公共预算承担。

在此制度安排中，“观测”不再仅仅是一种认知行动，而是一种制度行动。它要求主权国家让渡部分信息自主权，要求各国议会年复一年批准气象预算，要求成千上万的基层技术人员在午夜准时放飞探空气球。整套制度一旦建立，便反过来定义了气象知识的形态——那些未被纳入WMO标准的信息，无论在局部多么精准，在专业意义上都不构成“气象知识”。气象学的成功，是最干净地证明了“预测不是智识穿透混沌，而是制度在混沌上搭建脚手架”的活化石样本。

但恰恰是在这个最干净的成功样本中，埋藏着一颗关键的比较性种子。WMO的物质语法之所以能在超过半个世纪的时间里持续成功而不发生自噬，是因为它恰好完美地避开了三条机制的激活条件。第一，大气“不看回去”：大气不会因为被测量了温度而改变自己的气流路径，因此不存在被计量对象对计量语法本身的策略性规避。第二，公共管道未被私有化替代：虽然商业气象服务在近二十年显著增长，但它们的原始数据来源仍然深度依赖WMO的公共数据交换网络——不存在一个能以更低成本提供全球覆盖的、完全私有的替代性数据采集体系。第三，模型不构成交易对象的内战坐标：气象预测模型的输出虽然被广泛使用，但其使用者（农民、航空公司、应急管理部门）并不会通过修改自身行为来反身性地改变大气状态，从而形成“预测改变被预测对象”的自噬闭环。

这三种条件的干净不满足，使得气象学成为“制度化认知成功而未激活自饮性”的基准参照。但必须立即强调：这一参照地位的成立是历史的、有条件的，而非自然赋予的。它不意味着大气系统天然地“免疫”于自饮性；它只意味着在当前历史阶段，气象学所面对的特定对象与制度安排尚未耦合出自饮性机制。这一判断的逻辑强度恰恰来自其可证伪性：如果未来某一天，碳交易定价与气候风险金融化发展到足够深度，使得被预测的人类排放行为开始系统性地反向套利气候模型，那么气象学也可能部分地滑入自饮性——这恰恰说明自饮性是一种由特定制度条件所激活的机制，而非一种预先注定的本质属性。

接下来的三章，我们将从这一基准出发，逐一进入三个“成功且自饮”的领域——金融监管、公共信息管道、以及预测模型——以展示同样的制度成功，在不同的对象与条件之下，如何结算出截然相反的认知后果。

三、计量与套利的耦合：可读性如何喂养规避者

现代经济统计的可读性，是在二十世纪几代统计官僚、经济学家、议会预算委员会与国际组织之间，通过立法、行政命令与条约义务，一步步植入社会经济体的骨骼之中的。国民账户体系、货币供应量、通胀率、失业率、国际收支平衡表——这些范畴没有任何一个是“天然地就在那里”的数据。它们是从战后英美围绕重建国际经济秩序的制度之争中，被凯恩斯、理查德·斯通及其各国同行当作一套统一的会计语法锻造出来的。1953年联合国发布第一版国民账户体系（United Nations, 1953），接受马歇尔计划援助的欧洲国家被要求以此格式提交经济数据——这套语法在其诞生之初，就是国际政治博弈的产物。

问题在于，与大气不同，经济行动者会“看回去”。一旦“被清晰计量”意味着更高的税率、更严格的监管审查、更透明的竞争博弈，三个要件便被准时激活：动机（规避的收益）、能力（精确理解语法边界并设计规避工具的技术手段）、机会（制度缝隙提供的物理操作空间）。过去半个世纪的金融创新史，便是这套计量语法与其规避者之间持续咬合的编年史。二者之间的关系并非某种“本体论上预先注定的孪生对称”，而是在每一个具体的历史节点上，经由特定的法律文本、技术手段与市场条件，才被结算为“监管愈细、规避愈精”这一看似无解的耦合形态。

需要一次更精确的时序追踪来展示这一耦合如何在具体的制度纹理中被锻造。1988年，巴塞尔协议I引入了风险加权资产的概念——银行须对不同类型的资产持有不同比例的资本金。其中，商业贷款被赋予100%的风险权重，这意味着每发放一美元的商业贷款，银行必须保留8美分的资本金。这一条款的初衷是审慎监管，但其直接结果是：任何能将贷款“转换”为较低风险权重形式的金融工具，都能为银行释放出巨额资本金，从而显著提升其杠杆能力与利润空间。资产证券化正是在这一精确的激励机制下从边缘创新转变为主流工具：通过将一个贷款池打包为证券并出售给特殊目的载体（SPV），贷款从发起银行的资产负债表中消失，相应的资本金要求随之蒸发。而SPV本身的注册地——开曼群岛、泽西岛、特拉华州——其法律形态与巴塞尔文本中的“并表监管”条款之间，只隔了一层可以被合法穿透的帷幕。

此次创新并非监管者的疏漏，也不是投机者对规则的暴力破坏。这是在监管语言与法律语言之间天然存在的语义缝隙中，被精确设计出来的一种新的制度安排——它完全合法，但它所创造的那片金融活动却从原初计量语法的可读范围内脱逃了。帕兰（Palan, 2003）在《离岸世界》中详尽论证了离

岸金融并非“法外之地”，而是在各国法律主权的缝隙中精心设计的制度安排——每一个离岸实体都是在一个司法管辖区内完全合法、在另一个管辖区内有选择性不可见的复合构造。

1990年代，信用违约互换（CDS）被发明，允许机构持有某种资产的风险敞口却无需在会计上体现该资产本身。2000年代，影子银行体系——货币市场基金、对冲基金、特殊目的载体、结构性投资工具——迅速膨胀，形成了一个与受监管银行体系平行、却几乎不被同一套计量语法所捕获的金融宇宙。根据金融稳定理事会（FSB）的年度监测报告，全球非银行金融中介的资产规模在2010年代末至2020年代初持续增长，其绝对体量已超过全球GDP的总和，其中相当大比例的活动通过跨境司法套利绕过了单一国家的监管雷达。

这一演化过程并非一蹴而就，而是分阶段的、有断裂的、在每个节点上都存在着其他可能路径。1980年代的美国储蓄贷款协会危机催生了巴塞尔I；巴塞尔I在提升银行体系透明度的同时，逼出了证券化浪潮；2008年全球金融危机暴露了影子银行的系统性风险后，巴塞尔III试图堵漏（例如通过引入杠杆率和流动性覆盖率），但随即又催生了新的规避形态——如风险转移交易和合成型证券化。2004年巴塞尔II引入的内部评级法，本意是允许大型银行使用自己的风险模型以更精细地匹配风险与资本要求，但其实际效果之一，是为银行开辟了通过模型假设来系统性低估风险权重的新通道，从而进一步激励了“模型套利”。每一个回合的制度回应，在修复了一处漏洞的同时，都为下一轮规避提供了新的制度缝隙与新的工具——这不是静态的对称结构，而是一场持续漂移的军备竞赛。

正是在这种历史化的意义上，我们将此过程称为“耦合”而非“孪生”：它不是本体论层面预先注定的，而是在每一轮具体博弈中被结算出来的暂时稳定态，每一次结算都可能是另一个样子——只是迄今为止，每一次博弈的结果都指向了同一个方向。

这也解释了为什么“另类数据”方案——通过卫星估算超市车流量、通过信用卡账单模拟零售消费——在提供局部领先指标上偶有建树，却永远无法恢复整体的预测能力。这些方案所做的是，在默认相当一部分经济活动已经合法地滑脱了传统计量语法之后，试图在骨架外围加装高灵敏度的感应器，来间接追踪那些从主管道中滑出的血流。它所捕捉的，是滑脱者在移动时偶然溅出的几滴血——从未是滑脱本身。当整个金融体系的计量基础设施被制度性地、合法地规避时，对感应器精度的升级，就只是对系统性认知盲区的数字化代偿。

四、公共管道的内爆：信息基础设施的私有化如何瓦解公众认知

如果经济预测的认知耗损源于其计量语法在成功中喂养了规避者，那么社会与政治预测领域的认知耗损所遵循的，是同一逻辑在另一层维度上的显影。

气象学成功的核心机制，不仅在于有一套统一的物质语法，更在于有一个由各国公共预算共同维护的“公共管道”——一个将所有观测站数据汇聚、整理、分发至所有预测参与者的全球中继系统。同理，一个社会若要形成对自身“民意”的可读认知，也需要一种功能上等同的公共信息基础设施：一套能够将个体意见汇聚为集体公共表述的、由公共投资支撑的采样与分发系统。

这套公共管道的制度化，是二十世纪大众传播与代议制民主共同演化的物质产物：随机抽样的入户访问、地方报纸的社区新闻、公共广播的权威播报。它们的物质效果，是确保了几件对社会自我认知至关重要的事——不同阶层的人在相同的样本框中相遇；市政公告几乎同时抵达雇主与雇员；晚间

电视新闻提供一份关于“今天世界发生了什么”的相对共享的清单。沃尔特·李普曼（Lippmann, 1922）早在1922年就指出，“民意”从来不是像气温一样等待测量的自然事实，而是被一套有形的媒介基础设施所生产出来的集体表述。这套表述的真实性，依赖于信息采集与分发的公共管道保持广泛的、跨越社会阶层边界的覆盖能力。

过去二十年间，发生在物理层面的事实是：社交媒体平台和算法推荐系统，在信息分发的效率上彻底碾压了这套公共管道，并从内部取代了它。信息的分发不再主要经过公共投资支撑的中继神经——印刷厂、广播频段、公共电波——而是经由全球科技公司的私有服务器阵列。这一转变的性质，不应被误读为“自由市场对臃肿官僚体制的升级”，而应被如实描述为一种信息基础设施的私有化：如同供水系统被从市政管网整体置换为私有钻井公司的集成。

但是，将这一过程叙述为“锈蚀”或“外部侵蚀”是误导性的。公共管道的消逝，并非年久失修，而是被一种更高效的、私人化的信息分发方案从内部取代——这更接近于一种“内爆”：一座建筑不是被外力推倒，而是其内部承重结构在自己所支撑的重量下发生了不可逆的崩溃，而崩溃所释放的空间，被另一种承重逻辑即刻占据。

这一更高效的私人方案之所以能诞生并在技术上迅速成熟，恰恰是因为它被旧公共管道先前的成功所浇灌出来。二十世纪大众传播制度在“制造跨阶层公众”上的巨大成就，培育出了海量的、对信息多样化存在需求却被均质化供给所压制的用户群体——正是这些被旧体制喂饱但又未被完全满足的需求，构成了私人算法最初的市场激励信号。与此同时，同一套大众传播制度所催生的数字通信基础设施（光纤、协议、编码标准）和受过高等教育的工程师群体，为私人算法的诞生提供了所有的技术砖块。曾经服务于公共认知的原料——光纤网络、数据传输协议、内容格式标准——被私人平台重新编织进了一套完全不同的分发架构，这套架构优化的不再是“公众的共性”，而是“个体的差异”。

算法推荐的运作逻辑恰恰与“公共”的构成条件背反。它被设计出来，并非为了均质化信息以使公众形成重叠共识，而是为了最大化每个个体的驻留时长——其方法，就是将信息流切割、分筛，为每个用户单独重构一个刺激最大化的微环境。当每个人逐渐活在自己独一无二的私人信息茧房中时，“公共舆论”这个概念本身便遭遇了基础层面的瓦解：抽样调查的方法论基础——从一个共享共同信息输入的总体中随机抽取个体，并假设他们的意见可以被有意义地加总——在物理上被逐步掏空。

以2016年美国总统选举中民调机构的系统性误判为例。传统民调的抽样方法论，建立在随机拨打电话（包括座机与手机）并假设样本能代表登记选民总体的基础之上。然而，在几个关键摇摆州，特定选民群体——例如没有大学学位的乡村白人男性——已经大幅迁移至私人化的信息生态：他们在社交平台群组、加密通讯应用和定向播客中交流政治信息，对传统媒体和陌生电话采访的信任度与应答率急剧下降。民调机构通过统计加权来纠正样本偏差的努力，恰恰撞上了更根本的困难：加权所依据的总体参数本身，正是从被私化信息管道所撕裂的旧有社会范畴中推导出来的。当“拥有大学学位”或“城市居民”这类分组变量，已经不再能捕捉不同群体所居住的完全不同的信息宇宙时，加权的技术工具箱就变成了一座架在裂缝上的精密天平——它仍然在测量，但秤盘下面的地面已经分裂。这一失灵不是模型技术的偶然失误，而是一个更根本的信号：民调赖以运作的那套公共信息基础设施，正在被其所催生的私有替代品从内部瓦解。

这并不是一个关于“自由市场侵蚀公共领域”的简单道德叙事。这是一套制度化公共认知系统在其自身成功中所必然积聚的、最终被私人的、更高效的技术从其内部爆开的张力。曾经属于公共认知的原料——信息分发管道、数据格式、用户注意力——在旧制度的成功中被培育成熟，却被另一种制度逻辑所收割。公共管道的内爆，与金融监管层面的“监管吹胀套利”，遵循着同一条深层逻辑：成功的制度化秩序，在结构上持续地将自己转化为喂养对抗性系统的养料。

五、预测在成功中自噬：从操演现实到切断现实

即使物质语法未被规避、公共管道未曾内爆，仍有第三重机制内置于制度化认知的成功之中——而这一次，它不需要任何人“钻空子”。

唐纳德·麦肯齐（MacKenzie, 2006）对布莱克-斯科尔斯期权定价模型的制度史追踪，为观察这一机制提供了最干净的窗口。1973年该模型发表之初，模型所算出的理论价格与市场实际价格之间存在显著偏离。但随着模型被芝加哥期权交易所的交易员广泛采纳——被装入交易终端、风控系统、监管公式——市场价格在1980年代中期逐渐收敛于模型价格。这不是因为模型“发现”了既有的经济现实，而是因为模型被广泛装入了市场的基础设施，从而在物质层面表演出了它所应描述的现实。麦肯齐将此称为“有效的操演性”。

但操演性是一把双刃剑，其反面锋刃与正面锋刃由同一块铁打出。要理解这一转化机制，必须精确追踪其发生的微观过程。当模型最初被少数交易员使用时，它的主要功能是引导交易决策——交易员使用模型估算期权的理论价格，而后对比市场价格进行套利交易。在这个阶段，交易员的行为所针对的参照系，是“模型所描述的现实”——即基础资产价格的波动率本身。但当模型逐渐成为全市场的基础设施——被装入自动定价软件、风险管理模块、交易终端的默认参数——之后，反应对象发生了关键性的位移：交易员不再对“波动率本身”做出反应，而是对“模型输出的波动率数值”做出反应。一个交易员的计算不再从“这只股票的波动率是多少”出发，而是从“如果其他人都用布莱克-斯科尔斯公式，我该把自己的波动率参数设在多少才能在博弈中占优”出发。在那一刻，反应对象从外部现实滑向了模型自身，一个自我指涉的闭环就此咬合。

这一机制的极致形态出现在高频交易的微观世界中。当算法在毫秒级别所做的工作就是去预测其他算法的瞬间反应时，预测的终极参照坐标不再是劳动生产率或资本积累这类缓慢移动的实体地壳，而是预测机器之间的高速内战——每台机器都在用自己的预测去揣测、反制、先发制其他机器的预测，而那个外部参照系在这场内战中逐渐退为一个遥远的背景噪音。本文将此命名为自噬性操演：模型在被广泛采纳后，先是通过成功表演现实而实现有效的操演；随后，当它彻底成为市场基础设施、反应对象从现实位移至模型本身时，有效的操演性便在自己的成功中滑入一个自我指涉的闭环，逐渐切断其与外部参照系之间的信息神经。

必须在此处明确划出自噬性操演与两个最邻近理论传统的边界。其一，索罗斯（Soros, 1987）的“反身性”涉及认知与现实之间的普遍相互渗透——两者在一般状态下就互相影响。自噬性操演与此不同：它特指模型从“表演现实”到“取代现实成为反应对象”这一结构性位移。反身性不必预设这一位移；它允许的只是参与者的预期在边际上影响了他们所预期的对象，而无需切断模型与外部参照系之间的主干道。其二，卡隆（Callon, 1998）与费拉罗等人（Ferraro et al., 2005）所发展的“反向操演性”（counter-performativity），描述的是模型被广泛使用后反而使其所描述的现象

失效或消失——例如，一个预测股市崩盘的模型被广泛相信后，反而阻止了崩盘的发生。自噬性操演与反向操演性的关键区别在于：后者是操演效果在“方向”上的反转（模型的使用产生了与模型预测相反的结果），前者则是操演在“参照系”上的整体位移——模型的使用不必然产生反向效果，而是在自身的成功中逐渐把参照系从外部现实位移到模型自身，从而建造出一个与外部现实渐行渐远的闭环游戏。反向操演是“成功摧毁了自身”，而自噬性操演是“成功把自身变成了一个独立的世界”。

菲利普斯曲线的完整命运便是这一机制的宏观经济注脚。1958年，菲利普斯发表了那条稳定的工资变化率与失业率之间的此消彼长曲线后，它一度成为各国宏观政策制定的可靠罗盘。然而，在1960年代末至1970年代，当政策制定者、企业和工会都开始深刻相信并按此曲线行动时，企业的定价开始包含预期的通胀率，工会的谈判开始锚定预期的生活成本上涨，而预期通胀率的数值本身正是由包含菲利普斯曲线在内的那套宏观模型所产出的。在那一刻，行动者的反应对象不再是劳动力市场上工资与失业之间的实际此消彼长关系，而是模型所描绘的那条曲线——他们针对那条曲线调整行为，从而使那条曲线本身发生漂移。大滞胀这一历史事件不是证明菲利普斯曲线一开始就错了，而是它太对了、太成功了、太被相信了——被预测者广泛地内化进了自己的决策方程，从而系统性地改写了那条曲线得以存在的制度与行为条件。

气象学的物质语法之所以在长达半个多世纪里持续成功而未陷入自噬，正是因为它恰好避开了上述三重机制的全部激活条件：大气不看回去，故无规避者；其数据采集的公共管道没有被私有化替代——因为不存在一个能以更低成本提供全球覆盖的私人替代性数据采集体系；大气本身也不根据模型输出的数值调整自己的行为，故不存在“反应对象从大气位移至模型”的自噬触发点。这三重条件的干净不满足，使气象学成为“制度化认知成功而未激活自饮性”的基准参照。但正如第二章已经强调的，这一参照的地位是历史的、有条件的，而非自然授予的永久豁免——它恰恰通过对比，揭示出了在其他领域中，制度化认知的成功自噬是如何被其对象的结构性特征所精准激活的。

六、自饮性的统一内核：秩序的生产、流通与反馈

前面三章的论证铺开了三条并行且交叠的机制链，每一条都指向同一个方向：

- 计量语法的制度化成功，激活了被计量对象的规避动机、能力与机会，从而喂养出系统性的套利行为——并非“漏洞被投机者发现”，而是成功本身逼出了逃逸其自身可读性的反向生成链；
- 公共信息管道的制度化成功，培育了海量用户群体与技术原料，最终被更高效的私有算法从内部取代——并非“被外部腐蚀”，而是成功所积累的资源被另一种制度逻辑收割；
- 预测模型在制度化成功中，从操演现实滑向切断现实——反应对象从外部世界位移至模型自身，形成与参照系渐行渐远的自我指涉闭环。

这三条链并非彼此无关的三类“失灵原因”。它们共享同一个深层结构性内核，而将这个内核清晰地提炼出来，是令“自饮性”从一个现象层面的敏锐发现上升为一个具有解释力的中层理论的关键一步。

让我们将三条链并置在一个共同的分析框架下观察。在任何制度化认知体系的运行中，都同时存在着三个不可彼此还原的功能性环节：秩序的生产（一套物质语法、分类标准与计量规则的建立与强

制）、秩序的流通与物化（这套秩序在公共管道、信息基础设施和制度化实践中的传输、沉淀与日常运作）、以及行动者对秩序的反馈（被计量、被分类、被预测的对象，根据这套秩序所传递的信号与激励，调整自身的策略性行为，并将调整后的行为重新输入回这套秩序之中）。这三个环节不是时间上的先后阶段，而是任何制度化认知体系在持续运转中所必需的功能维度——生产端的稳定、流通端的畅通、反馈端的校准，三者缺一不可。

金融监管的自饮性链条，在生产端（巴塞尔协议制定的风险权重分类）与反馈端（银行据此设计证券化工具以规避分类）之间的咬合最为突出。公共管道的内爆链条，则主要显影于流通端：旧有的大众传播管道在成功中将信息分发的基础设施、用户注意力与技术标准作为“公共品”生产出来，而这些流通端所积累的资源，恰好构成了私人算法平台从内部取代其分发功能的全部原料。预测模型的自噬链条，则集中发生在反馈端：当模型成为交易的基础设施后，反馈环节的参照系从外部现实位移至模型自身，反馈的校准功能逐步失灵，形成自我喂养的闭环。

这三条链因此并非孤立的类比，而是同一自饮性机制在制度化认知体系的三个功能位点上的不同显影。每一种显影的具体形态，取决于该领域在这一特定历史阶段，究竟是哪一个功能环节上的“成功-反馈”循环率先被激活并发生了结构性的偏离。在金融监管领域，是生产端对反馈端的“逼迫-回应”循环率先咬合；在公共信息领域，是流通端的成功积累为流通端的私有化提供了全部前置条件；在预测模型领域，是反馈端的过度成功使其切断了与生产端的校准神经。由此，我们可以将自饮性的统一内核精确地表述为：

制度化认知的成功运行，必然同时增强其在秩序生产端的技术能力与秩序流通端的物质积累；然而，当这一成功深度嵌入被认知对象的激励结构之后，反馈环节便会从“校准认知”转向“策略性利用认知”；这一转向会持续将生产端的分类语法、流通端的公共管道以及反馈端的参照系，转化为喂养反向系统的养料，从而系统性地耗损该制度体系对其对象的认知能力。

正是这个内核，将自饮性与古德哈特定律、免疫自噬论、卢曼的自创生以及反身性/操演性等相邻概念清楚地区分开来。本章已经给出了机制的统合论证；下一章，我们将把自饮性直接置于这四个最强大的理论对手面前，通过严格的概念边界划定，证明它并非对任何一个已有理论的修辞换皮，而是切开了它们各自仅能握住一侧面、却整体装不下的那条辨别之缝。

七、不可还原的边界：与四个最邻近理论的对话

本章是本文概念合法性的核心环节。在此，将自饮性依次置于以下四组最邻近且最强大的理论传统之下，进行代理坍缩分析与不可还原边界的严格划定：古德哈特定律（Goodhart's Law）、埃斯波西托的免疫自噬论（immunization/autoimmunity）、卢曼的系统自创生（autopoiesis）、以及反身性与操演性文献（reflexivity / performativity）。若自饮性可以被其中任何一个理论加上一个修饰词而无损重述，则本文的核心贡献便不成立。

7.1 古德哈特定律：“指标即目标”之后

古德哈特定律（Goodhart, 1984）——常被简化为“当一个指标成为政策目标，它就不再是一个好的指标”——无疑是讨论“制度化可读性喂养规避”的奠基性传统。该定律精准地揭示了一旦某个

统计量被赋予决策权重，被观察者便会策略性地操纵该统计量，从而导致该指标与其本应收敛的现实之间脱节。

自饮性与古德哈特定律的核心边界在于机制的层级与反向系统对母体的本体论依赖。古德哈特定律描述的是一种信号失真：指标被操纵后，其所传递的信息质量下降。它的默认前提是：仍存在一个外部的、独立的“真实”经济状态，只是指标不再能忠实地反映它。但自饮性所描述的不仅是信号失真，更是制度基础设施的逆转。在古德哈特定律中，被操纵的对象是“指标本身”；在自饮性中，被喂养壮大的是一整套平行于、并依靠汲取母体制度而存在的反向物质系统——在金融案例中，这不仅意味着“货币供应量M2不再可靠”，而是意味着整个影子银行体系作为一套完整的制度安排，依靠持续吸取受监管银行体系的资本、人才与法律工具而膨胀。前者是信息层面的脱耦，后者是制度层面的逆生。

一个2×2判别表可以澄清这一边界。设定两个维度：纵轴为“是否存在一套物质性的反向制度系统（而非仅仅是数据操纵行为）”，横轴为“反向系统是否以母体制度的维持为其存续的必要条件（而非独立运作）”。古德哈特定律覆盖了“无独立反向系统”的区域——行动者操纵指标，但不必然发展出一套完整的、自我维持的替代性制度基础设施。自饮性则覆盖了“存在独立反向制度系统”且“该系统以持续汲取母体制度为生”的区域——影子银行并非简单“扭曲”了商业银行的会计科目，而是发展出了平行于银行体系的、自身拥有特殊目的载体、法律契约模板与二级市场流动性的完整替代架构，而这一架构的持续运转，恰恰需要受监管银行体系的持续存在作为其原料来源（风险资产、合规外壳、支付结算终端）。正因为如此，自饮性所导致的认知耗损，远比古德哈特定律所预测的“指标失效”更为深层和难以逆转：要恢复可读性，不是更换一个指标，而是拆除一整套已经制度化、并持续吸取旧制度养分而活的物质替代系统。

7.2 免疫自噬：纳入他者，还是被他者汲取？

埃斯波西托（Esposito, 2002, 2008）的免疫理论为理解“系统如何通过纳入一部分威胁物来保全自身”提供了极为精致的分析框架。在免疫逻辑中，共同体通过接种（纳入弱化的他者）来产生抗体，但这种纳入在某些条件下会滑向免疫性自噬——系统错误地攻击自身健康的组织。

自饮性与免疫自噬的核心区别，不在于“攻击”与“摄食”的修辞差异，而在于发生路径的方向与系统的本体论状态。在免疫自噬中，系统仍是主动方——是免疫系统自身的识别机制发生了紊乱，导致了错误攻击。在自饮性中，系统既非主动也非紊乱；它只是在它最成功的一刻，在物质层面生产出了一套以自己为食的新系统，而母体制度本身的识别与执行机制仍在照常运转——它并没有将自己识别为攻击目标。证券化机制并未使巴塞尔协议的资本充足率计算公式发生紊乱；公式依然在精确运行，它只是在公式所规范的对象外部，长出了一整套公式触及不到的替代性金融架构。母体制度没有“攻击自己”，它只是被持续地汲取，直至衰竭。

因此，自饮性提供的是一个免疫理论无法整体装下的辨别维度：反向系统的诞生，不是因为母体系统的识别机制出了故障，而是因为母体系统的分类语法过于成功，以至于在它所分类的对象之中，孕育出了一个在语法之外的、却又完全以该语法为生存依据的新系统。免疫自噬是“系统的自我攻击”，自饮性是“系统在成功中分娩出以自己为食的寄生体”。

7.3 卢曼的自创生：运作封闭能否容纳“被汲取”？

卢曼（Luhmann, 1984）的系统理论将社会系统界定为运作上封闭的自创生系统——它们通过自我指涉的沟通不断地再生产构成自身的元素，与此同时对外部世界（环境）保持结构性盲目。这一框架已经内在地包含了“系统的成功运作必然意味着对某些外部刺激的麻木”。

然而，自饮性与卢曼式自创生的根本差异在于环境与系统之间关系的拓扑性质。在卢曼的架构中，环境只是系统运作的“不可达的彼岸”——环境刺激可以扰动系统，但永远无法进入系统；系统从环境中“汲取”的，只能是它自身运作所需的能量或信息形式，而不会威胁到其自创生的同一性。但自饮性所描述的现象，恰恰推翻了这个拓扑前提：反向系统并非外部环境，它就是从母体系统自身的成功中分化出来，并且持续从母体系统内部汲取制度性的养分——资本、合法性外壳、专业人才、法律技术——来维持自身。它不是“彼岸的噪声”，它是同一制度生态中、由同一套成功所催生的“此岸的竞争者”。卢曼的“自创生”关注的是一个系统如何在环境中维持自身边界；自饮性关注的则是，一个系统的成功可以如何在自身内部催生出一个以自己为食的新系统，而这两个系统共享同一个制度内环境。

用一个区分来凝结这一差异：卢曼的问题是“系统如何在其运作封闭中实现自我维持”，自饮性的问题是“制度化认知系统的成功，如何在结构上必然导致其中一部分养分被内部衍生的替代系统持续抽走”。后者不是运作封闭，而是养料分流。

7.4 反身性与操演性：位移，而非仅仅反转

最后的边界划定已在第五章做了预埋，此处仅以最精简的形式呈现以便形成完整的边界系统。自噬性操演与索罗斯反身性（Soros, 1987）的区别在于：反身性是普遍的互渗，自噬性操演是从操演现实到切断现实的特定位移。自噬性操演与反向操演性（Callon, 1998; Ferraro et al., 2005）的区别在于：反向操演性是“成功产生了反向效果”，自噬性操演是“成功把模型本身变成了一个自足的、与外部参照系渐行渐远的博弈世界”——前者还保留着对外部参照系的否决（因为效果发生了反转），后者则已经让参照系退为背景噪音。

综上所述，自饮性不可被还原为上述四个理论中任何一个的变体，因为它切开了它们各自均无法整体装下的辨别面：反向系统不是母体的错误、扭曲、反转或外部噪声，而是母体在其自身成功中所结构性地分娩出来的、以持续汲取母体制度养料为存续方式的一套物质性替代架构。这一辨别面的识别，使得自饮性不再是“制度悖论”的又一个同义词，而是对一种特定类型的制度化认知耗损的首次单独指认。

八、可裁决性：自饮性在何种条件下成立？

一个新概念若要脱离不可证伪的修辞层面，必须给出可裁决的经验判据：在何种可观察的条件下，自饮性假说会做出与竞争理论不同的预测，且该预测可以被经验世界所检验？

基于前文的机制统合，自饮性机制的激活依赖于三个条件的共同满足：第一，反馈端的策略化——制度化的认知语法深度嵌入了被认知对象的激励结构，使得精确理解并规避该语法成为一种可带来显著收益的策略性行为；第二，替代性制度的可建设性——存在制度缝隙或技术条件，允许行动者

围绕母体语法构筑一套平行的、替代性的物质系统，而不仅仅是个体层面的零星规避；第三，母体养料的持续供给——反向系统所依赖的制度资源（合法性外壳、人才、资本、信息管道、技术标准）持续由母体制度的成功运行所生产与维护，因而反向系统有持续的动力与能力去汲取而非摧毁母体。

这三项条件构成一个“激活函数”。当它们同时具备时，自饮性假说预测制度化认知的成功将导致系统性的认知耗损；当任一条件缺失时，成功则可以相对稳定地维持。

这一框架能够产生一组可裁决的预测，并将自饮性与古德哈特定律在经验层面区分开来。以金融监管为例：古德哈特定律预测，当某个风险指标（如风险加权资产比率）成为监管目标后，该指标与真实风险之间的相关性将下降。自饮性假说则进一步预测，（1）围绕该指标将出现一套制度化的替代性金融架构（如影子银行体系），其规模与监管语法的成功程度（覆盖广度与执行力度）呈正相关；

（2）该替代架构的存续并不依赖于突破或摧毁监管语法，反而依赖于监管语法的持续运作——因为监管语法正是其“养料”（如提供可被证券化的合规贷款、维持母体银行体系作为最后的流动性后盾）；（3）单纯更换监管指标（如从风险加权资产转向杠杆率）无法根除替代架构，只会促使其发现新的规避节点，因为在完全拆除母体制度本身之前，养料源头仍在。

这三个预测在经验上可被长期追踪。已有的国际金融监管史（从巴塞尔I到III的演化）恰好构成一组自然实验：每一轮监管改革都在初期改善了风险指标的信号质量，但随后均催生了更大规模的、制度化的影子银行创新，而这些创新在监管指标被替换后并未消失，只是转移了管道。这与自饮性假说一致，而与“单纯指标失效”的古德哈特框架相悖——后者无法解释替代架构的制度化与持续自汲取现象。

在公共信息管道领域，自饮性假说同样给出了可裁决的预测：当传统民调方法论失灵时，增加样本量或引入更复杂的统计加权技术将无法修复其根本问题，因为导致失灵的原因不是抽样误差，而是目标总体本身在私有信息管道基础上的结构性分裂。这一预测可以被跨国比较检验——在公共广播制度相对坚韧、社交媒体渗透率较低、信息基础设施私有化程度较轻的国家，民调准确率应显著高于那些公共管道已深度内爆的国家。这同样是一个可被经验数据核实或证伪的判别格。

这些经验判据，使得自饮性不仅仅是一个精巧的概念构型，而是一套可供学术共同体检验、修正甚至证伪的分析框架。而给出其可能被证伪的条件，正是学术诚信的组成部分，也是下一章回应“效率论”反方时所依托的基石。

九、最强反驳与变异案例：为什么这不是决定论？

在建立了自饮性的概念并给出可裁决判据之后，必须直面两个最有力的反驳，否则本文的论证便会滑向一种决定论的、只能自我证成的“历史哲学”。

9.1 效率论反驳：这是高效创新，而非认知耗损

第一个最强反方来自新制度经济学的效率传统。这一传统可以有力地论证：影子银行的兴起，不是制度认知的“自饮”，而是金融市场在监管约束下的高效制度创新；它通过风险转移提升了整体资本配置效率，将风险从受到严格资本约束的银行体系转移至更具风险承载能力的非银行投资者手中。同理，社交媒体对公共管道的取代，不是信息基础设施的“内爆”，而是信息分发的帕累托改进——

它以更低的边际成本满足了异质性的公众需求，将信息的“计划经济”更新为“市场经济”。按照这种逻辑，自饮性不过是将成功的制度变迁在事后病理化——一切非预期的制度演化，都可以被重新描述为“自饮”，从而使这个概念失去独立的解释力。

对此反驳的回应必须认清其合理内核并加以承认，进而在关键处划出边界。影子银行确实提升了某些维度的金融效率，社交媒体也确实满足了被大众媒体均质化所压制的多样化信息需求。效率论错误的，不是这些效率收益的描述，而是它对于替代系统与母体系统之间关系的定性。影子银行的效率提升，不是两个独立系统之间的自由市场竞争结果，而是一个系统持续汲取另一个系统的制度性公共品而实现的寄生性效率。特殊目的载体的法律外壳、做市商提供的二级市场流动性、中央银行的最后贷款人功能、以及“太大而不能倒”所隐含的隐性担保——这些都是母体公共制度体系在其成功中所生产与维持的公共品。影子银行作为替代系统，并未承担这些公共品的再生产费用，它所做的只是精确地在法律缝隙中吸取它们。因此，它所实现的效率收益，不能被视作纯粹的市场效率，而是一种依赖于母体制度持续存在的不对称效率。一旦母体制度因被过度汲取而丧失了提供这些公共品的能力（例如，当持续性公共救助引发政治反弹并导致制度重构时），替代系统的效率也将随之瓦解——这正是2008年危机所暴露的节点。效率论证遮蔽了这一养料依赖的结构性事实。

同理，社交媒体平台的信息分发效率，同样深度依赖于公共部门所建设的物理互联网基础设施、资助的基础研究（如早期算法与数据压缩技术）、以及公共教育体系所生产的海量内容创作者。平台将这套公共基础设施所产生的正外部性内部化为企业利润，而并未承担这些基础设施的更新成本或对其信息外部性（例如公共认知撕裂）进行内部化。这不是帕累托改进，这是公共品被收割。

因此，效率论并未驳倒自饮性；它只是揭示了一旦我们采用一个不同的、将制度养料视为公共品的成本核算框架，自饮性假说所描述的成本便浮现出来。效率论是自饮性在同一枚硬币的另一侧投下的影子：它看到的，正是反向系统在汲取成功后的成长效率，而没有看到这成长是建立在对母体养料的不可逆消耗之上。

9.2 能动性反驳：这是策略选择，而非结构必然

第二个强反驳来自能动性理论：金融工程师和算法工程师不是被“制度语法逼迫”出规避策略的被动客体，而是在特定激励结构下做出理性选择的主动行动者。将规避行为描述为“被成功的计量语法逼出”，有将行动者的能动性还原为制度结构的决定论产物的嫌疑。进而言之，这会导致一种历史终结论：每一次制度修补都是徒劳的，因为自饮性会必然地再次发生。

这一反驳揭示了自饮性论证中一个必须被认真回应的危险倾向。在回应中，必须清楚地区分结构性的压力与决定性的必然。自饮性分析论证的是：当一套制度化认知深度嵌入被认知对象的激励结构、当制度缝隙为替代架构提供了可操作的建设空间、且母体制度持续为替代架构提供养料时，就会有强大的、持续的结构性的压力，将行动者的能动性引导向汲取母体养料的方向。这并不意味着每一个具体行动者都被剥夺了能动性——恰恰相反，他们正是在进行着理性的、创造性的、在给定约束下最优的策略设计。自饮性所要解释的，不是个体的微观动机，而是集体行动的宏观方向为何会系统性地指向某一端。这一方向性的锁定，是结构所强制的，但不是每一次具体行动所预先注定的。因此，自饮性描述的是结构化趋势，而非历史决定论。

9.3 变异案例：成功而未被自饮的子系统

论证机制的条件性而非必然性的最有力方式，是提供变异案例——在同一个宏观领域中，某些子系统同样制度性地成功了，却未激活自饮性，且这种“未激活”可以被自饮性的激活函数所解释。

以金融领域为例。并非所有成功的金融监管都催生了同样规模的替代架构。一个显著的对比存在于支付系统领域。全球主要经济体的批发支付系统（如美国的Fedwire、欧洲的TARGET2）均处于中央银行的直接运作与严格监管之下，其制度化的可读性极高——每一笔批发支付的金额、对手方与时间戳均被实时计量与记录。然而，支付系统并未像银行业资本监管那样催生出庞大的影子替代架构。原因恰在于自饮性激活函数的缺失：支付系统的“被计量”并未显著增加被计量者的成本——结算效率恰恰是参与者所追求的公共品；替代性支付架构的建设缺乏制度缝隙——中央银行的垄断地位与实时全额结算的技术架构使得建设平行的批发清算系统在经济上不可行；且反向系统若试图汲取支付系统养料（如利用央行货币信用），将立即触发母体系统最强的监管干预。这一个案中，自饮性激活函数的第一项条件（反馈端的策略化）便不满足，因而成功没有为反向系统提供动机。

另一个更具张力的变异案例出现在某些国家的地方自治财政体系中。某些北欧国家的市政财政高度透明，所有支出细目均在公开可查的数据库中实时公布。这种透明性制度的成功，并未催生大规模的市政规避产业，因为“被清晰计量”在这里所引发的，是地方选民的政治问责而非更高税率——换言之，透明性所改变的激励结构是政治声誉而非经济成本。当被计量所激活的反向压力指向的是政治竞争而非资本逃逸时，替代性制度建设的动机与能力都被显著削弱，自饮性的激活链条便在反馈端的第一环被中断。

这些变异案例并非自饮性的反证；恰恰相反，它们通过展示自饮性机制在条件不满足时的“未发生”，印证了该机制是有条件、可阻断的，而非一项可被随意外推的历史定律。它们为制度设计提供了一线并非怀旧的可能性：如果要阻断自饮性，工作不是退回到没有制度的“原初自然状态”，而是重新设计被认知对象的激励结构，使得“被清晰计量”所带来的反馈，不是规避的最高收益选项。

十、结论：在自饮性中追问新的认知组织方式

本文从预测失灵这一可观察的现象出发，拒绝主流解释所共享的“智识穿透混沌”预设，推进了一个硬转向：制度化认知的成功，不是认知耗损的“意外副作用”，而是其自身运行逻辑在特定条件下必然结算出的结构性反向生成物。这一反向生成物，本文将命名为自饮性——制度化认知在成功中将自身赖以成立的可读性语法、公共管道与参照系，转化为喂养反向制度的养料。

在论证中，本文首先展示了现代气象学如何作为一种制度化的认知体系在其成功中维持了可读性而不自噬，并以此作为对比，揭示出其在特定历史条件下所依赖的不可复制的免疫条件。随后，本文通过金融监管史、公共信息管道的私有化变迁、以及预测模型的自噬操演，逐一追踪了自饮性在秩序的生产、流通与反馈三个功能维度的具体显影，并将三条机制链统合于同一个结构性内核之上。通过与古德哈特定律、免疫自噬论、卢曼式自创生以及反身性/操演性文献的严格边界划定，本文进一步论证了自饮性并非对已有理论的修辞换皮，而是切开了这些理论各自只握有一侧的那条辨别之缝；并通过一系列可裁决的经验预测与变异案例，展示了机制的条件性而非历史决定论。

本文的结论并不提供一套关于“如何修复预测失灵”的操作手册。相反，自饮性分析所揭示的深层困境恰恰在于：任何试图“修复”旧有制度化认知体系的努力，一旦成功，就会重新进入自饮性的反馈环——你的修补越是成功，你对它的依赖越深，它喂养出的替代系统就可能越庞大。因此，真正需要追问的问题不是“如何恢复旧制度的视力”，而是：当旧有的制度化认知体系在结构性自饮中持续沉降，那些在不同认知组织方式之中的微弱尝试——分布式自治组织的透明治理、开源信息生态、公民科学网络——究竟是在旧土壤中萌发的新组织原则的雏形，还是仅仅自饮性在更高阶维度上的新猎物？回答这一问题所需要的，不是另一个更精致的诊断，而是对这些新的制度实验进行同样冷静的发生学追踪与分析。

最后，必须诚实地给出自饮性概念的证伪条件。自饮性假说的核心预测是：在激活函数全部满足的领域，制度化认知的成功将导致系统性的、不可通过单纯更换指标或升级模型而逆转的认知耗损。若未来在多个满足激活条件的制度领域中，出现制度化的替代系统在母体语法持续成功的情况下主动衰退或被市场内生淘汰的历史案例，且这种衰退不是由于母体制度的强力压制或技术替代，而是由于反向系统自身的动力衰竭——那么自饮性的“成功即养料”的核心机制便受到了实质性的挑战。这一证伪条件并非不可想象；或许在某个特定技术条件下，高度透明的分布式账本技术使得规避行为本身也变得完全可读，从而同时击穿了自饮性所依赖的“制度缝隙”条件。然而在那之前，自饮性作为对制度化认知成功与其认知耗损之间结构性关联的首次单独指认，将为我们审视当代社会的预测失灵、制度失效与技术迷思，提供一条不可忽视的分析轴线。

附论·最近邻切割：自饮性不是古德哈特定律、不是操演性、不是共有池悲剧

作者在关键词里坦诚列出了「古德哈特定律」与「操演性」——这份诚实值得敬重，但也把切割的责任摆到了桌面：一个主动承认自己站在两个成熟定律旁边的概念，必须证明它多切出了什么，否则就是这两条定律的合写。

第一个邻居是古德哈特定律（当一个指标成为目标，它就不再是好指标）。自饮性看起来是古德哈特定律的推广。分离线是决定性的：古德哈特定律讲指标与目标的关系（度量一旦被当作目标就失真），它的损耗发生在度量的有效性上；自饮性讲的是成功本身消耗其成功所依赖的公共认知基底——不是「指标失真」，而是「制度化认知每成功运转一次，就消耗掉一点它赖以成功的、由多方共享的、不可再生的认知共同基础」。古德哈特定律里没有「被消耗的公共基底」这个存量项；自饮性的全部要害就在这个存量项上——它像一口井，每次成功的取水都在降低水位，而水位不会自动回升。可裁决分离：古德哈特式失真可由「多指标、动态换指标」来缓解（不让任一指标固化为目标）；自饮性不能靠换指标缓解，因为被消耗的不是某个指标，是所有指标共同依赖的那个公共认知基底——换指标只是换了个吸管，接着喝同一杯正在见底的水。

第二个邻居是操演性（同前，麦肯齐的模型塑造市场）。自饮性与操演性的血缘：都讲预测/模型改变世界。切开：操演性到「模型造出了符合模型的世界」为止，它可以是稳定的成功（市场持久地像模型）；自饮性专讲这个成功不可持续的那一维——模型塑造世界所依赖的「世界原有的、模型未曾编码的丰富性」被模型的成功运转持续抽干，直到模型面对一个已被自己抽干了异质性的世界，反而丧失预测力。操演性是「符号成功塑造现实」，自饮性是「这种塑造以消耗现实的不可编码剩余为

代价、并终将因这剩余耗尽而失效」。可裁决：自饮性预测，一个被单一模型高度操演化的领域，其「模型未覆盖的异质性」（可用预测残差的结构、局外行动者的比例等操作化）随时间单调下降，且下降到临界点后模型预测力骤降。这条曲线操演性理论不预测，自饮性专门预测。

第三个邻居是哈丁的「公地悲剧」（tragedy of the commons）。自饮性说「成功消耗公共基底」，听起来就是认知版的公地悲剧。本文的分离动作：公地悲剧的机制是多个自利主体对共享资源的过度索取（每个人理性地多取一点，加总起来毁掉公地），它是一个协调失败/激励错配问题，原则上可由产权界定或治理规则解决（奥斯特罗姆已证明公地未必悲剧）；自饮性的机制不是多主体的过度索取，而是单个成功的制度化认知系统在正常、正确、无人越界的运转中就在消耗公共基底——它不需要「过度」，正常运转即耗损，因而不能靠界定产权或治理规则解决（没有人在滥用，制度在正确使用中自我掏空）。这是自饮性最不可还原的一点，也是它可裁决的地方：若一个制度化认知系统的基底消耗可由「更好的治理规则/产权安排」止住，则它是公地悲剧，自饮性的诊断多余；本文预测止不住——因为消耗内在于成功运转本身，而非源于治理缺位。

附论·可裁决预测：自饮不是玄学，是可测的水位下降

「自我耗损」这类命题极易变成不可证伪的悲观修辞。本文把它钉在三条可测的线上。

预测一（异质性单调下降）：一个被单一预测/评估框架高度制度化的领域，其「框架未编码的异质性」（预测残差的结构复杂度、不遵循主流框架的局外行动者比例等）随该框架的成功运转单调下降。若异质性不降或上升，则「成功消耗公共基底」被证伪。

预测二（治理不能止损）：改善治理规则或界定认知产权，不能止住制度化认知系统对公共基底的消耗，只要系统仍在正常成功运转。若更好的治理即止损，则自饮性塌回公地悲剧，无需新概念。

预测三（成功一脆弱耦合）：制度化认知系统越成功（预测越准、越被普遍采纳），其在基底耗尽临界点后的预测力崩塌越突然、越剧烈。若最成功的系统崩塌得最平缓，则自饮性对「成功即自我掏空」的刻画错误。

参考文献

- [1] Goodhart, C. A. E. (1975). Problems of monetary management: The U.K. experience. Papers in Monetary Economics, Reserve Bank of Australia.
- [2] Ostrom, E. (1990). Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge: Cambridge University Press.
- [3] Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. American Journal of Sociology, 113(1), 1-40.
- [4] MacKenzie, D. (2006). An Engine, Not a Camera: How Financial Models Shape Markets. Cambridge, MA: MIT Press.

关于本文的创新智商标注

本文在原始投稿基础上，由编辑依王德生《SDE本体论》与 SDE 创新智商评估规程做了「最近邻切割」与「可裁决预测」两节加固，意在抬高其不可还原性（I）与可证伪性（F）两维。依评分规程铁律，任何人不得为自己参与修改的文本认证分数，故本文页面所涉创新智商为**修改设计目标（135–139 区间）**，非认证分；认证分须由独立评分者盖出处另行给出。