

度量同构异质体：当精算的尺子量出了一种「新医生」

指标没有造假——费用真的降了、住院日真的短了；被换掉的是生产这些数字的那个人：外形同构，内部质地置换。

蔡彦 著 · 之28 · 卫生经济 · 度量制度 · 发表于2026年7月31日 · 约 2.0 万字 · 预备学员

摘要

按病种付费（DRG）在中国落地十年，一个反直觉的窘境反复上演：次均费用与平均住院日在纸面上稳定下滑，医生与患者的不安却同步上升。本文试图论证，这并非控费力度不足或执行走样，而是所有依赖单一简化量尺来治理高度复杂专业实践的改革，共享一个隐秘的底层预设——“善意前提”——在医疗这个场域中必然失效。这个前提预设行动者的博弈始终留在制度划定的度量语言之内；但现实中，当一套费用度量标准被嵌入一个持有强大内部判断体系的专业共同体时，它不会简单地“扭曲”既有的行为，而是会催生出一个与原逻辑外表高度相似、内部质地却被彻底置换的“同构异质体”。本文以“度量同构异质体”为核心命名，系统界定其概念边界、发生机制，并通过提出可观测的判别标准，将其与制度性机会主义、博弈均衡和惯习理论三个最接近的理论对手划清不可还原的界限。文章先以重症肺炎病房的类型化叙事完整演示其演化过程，随后在与“结果主义辩护”和“专业主义自我辩护”两个最强反驳的正面对抗中，论证此种认知置换的不可逆侵蚀。本文最终判断，走出这种困局，不能靠更精细的精算，而必须引入一种独立的、不可通约于费用的健康价值度量，并刻意制造两种度量之间无法被调和的持久矛盾——经两种度量共同校准的临床实践与制度安排，只能从这种矛盾中被逼出来，而不可能由任何预先设计的完美方案平稳推出。

关键词

度量同构异质体；DRG；善意前提；单一度量暴政；价值度量；认知熔合

一、引言：当精算的尺子量出了一种“新医生”

按病种付费（DRG）被引入中国医疗体系时，携带一个逻辑自洽的承诺：让医院自发控制成本，终结按项目付费时代的过度医疗。然而落地十年后，推开任何一家DRG试点医院的财务科室，听到的往往不是“控费成功”的经验，而是一种更微妙的窘迫。重症患者被委婉地建议转往未实施DRG的上级医院；一次本该完整的治疗过程被拆成两次入院；病历上的诊断编码被精心“升级”，让病种分组从低权重区滑入高权重区；必要的康复项目在临床路径中被逐步精简。一边是纸面上持续下降的次均费用，另一边是从医生群体到患者群体的广泛不安。

这个窘境通常被归咎为“执行走样”或“监管不严”，仿佛只要分组拆得更细、权重算得更准、飞检来得更勤，问题就会收敛。但本文试图论证，问题的根源远比这更深。DRG的设计——以及它所代表的、用一把简化量尺来度量高度复杂专业实践的所有治理技术——共享一个从未被说破的底层预设。这个预设可以这样表述：只要制度设计得足够精确、激励足够合理，行动者就会朝着制度设定的目标调整自己的行为；即使调整过程中会有摩擦、偏差甚至博弈，这些博弈仍然会留在制度所划定的度量语言及其逻辑框架之内——行动者至多在这些度量标准上造假、夸大、钻漏洞，而不会发展出一套不可被该框架识别的替代性判断体系。换言之，制度制定者默认，被治理者的回应策略，无论多狡猾，都仍然是用制度自己的语言——费用、编码、权重、分值——来写就的。制度因此始终握有辨识、分类和纠正这些行为的最终主权。

本文称这个预设为“善意前提”（good-faith premise）[1]。它不是在说行动者道德上善良，而是在说一个更隐秘的认知预设：制度与其接受者之间共享一套度量语言，后者即便博弈，也“善意地”留在这套语言的引力场之内。制度制定者预想中的最坏情况，是行动者在费用和编码上做手脚。他们从未预见到另一种可能：行动者不是在这些度量标准上“造假”，而是整个人的认知结构——怎样看一个患者、怎样写一份病历、怎样判断什么是一个“值得收治”的病人——被这些度量标准在深层悄悄改写。这种改写一旦完成，行动者自己甚至不觉得这是在博弈；他觉得这就是正确的临床操作、就是一份完整病历该有的写法、就是一个理性医生该做

的收治判断。当“违规”与“合规”在同一个行为中无法被分离时，制度不是被绕过去了，而是被彻底悬空了。

正是这个从未被预见的后果，构成了一切“单一度量暴政”最深层的破坏力：它不只在制度的边缘地带催生机会主义行为，它更在制度的核心地带催生一种与原有专业实践外表高度相似、内部认知质地却被彻底置换的“新医生”。本文把这个被度量标准反向塑造出来的、与原有专业人格“同构异质”的产物，命名为度量同构异质体（metric-induced homomorphic heterobody）。

接下来的论证分六步展开。第一步，进入一间重症肺炎病房类型化叙事，完整展示度量同构异质体如何从零散的生存策略一步步凝聚为新常识。第二步，拆解“善意前提”如何导致制度在落地时必然裂开。第三步，给出“度量同构异质体”的完整定义、内部构造与发生机制。第四步，将它置于与制度性机会主义、博弈均衡和惯习理论这三个最接近的理论对手面前，厘定其不可还原的理论增量。第五步，在与“结果主义辩护”和“专业主义自我辩护”两个最强反驳的正面对抗中，论证此种认知置换的不可逆侵蚀。第六步，论证出路：真正的破局不在更精细的精算，而在刻意制造两种不可通约的度量之间的持久矛盾。

二、一间重症肺炎病房的解剖：同构异质体如何在一线“长出来”

要真正理解一个度量同构异质体如何在一个活生生的科室里被发生出来，不能只停留在概念辨析，必须进入一个具体现场。[2]本节以重症肺炎这个在DRG支付体系下典型的“不划算”病种为例，构建一个完整的发生叙事。叙事中的人物、科室、对话均已匿名化或合成处理，其功能不是实证某一个特定案例，而是展示一种发生学的生成过程：当DRG的定额压力落到一个具体病种上时，博弈行为如何沿着临床决策的缝隙一步步渗透，逐步精炼，并最终固化为科室“常识”与“正确做法”的全部过程。读者可以将其当作一种“如果典型条件齐备，这个过程将以大致如此的方式展开”的机制演示。

2.1 第一周：定额来了，科室的恐慌

假设我们走进某三甲医院呼吸与危重症医学科。DRG支付的第一个月，月度分析会上，一份财务报表被投影在墙上：重症肺炎组定额支付标准为1.6万元。但过去三个月收治的十六例重症肺炎患者——大多是高龄，合并多种慢性基础病，有的需要无创呼吸机——实际平均成本是2.4万元。一例平均亏损八千元。科室当月奖金池缩水了将近三分之一。

主任沉默了一会儿，说：“我们不能不收肺炎病人。但也不能一直这么亏下去。大家想想办法，看看哪些地方可以省。”

这第一周里，医生们做的调整都是完全合理的、在医学上也有依据的：能用集采目录里的廉价抗生素就不用原研药；能在门诊做的检查就不占用住院床日；符合出院标准的患者不再像以前那样多观察一天。这些调整在本意上，是临床路径的正向优化。但一个变化悄悄地发生了：每一次查房，医生比以往多问了一句——“这个检查真的必须住院期间做吗？”——这句追问本身是合理的，但它在认知上已经种下了一个悄然的位移：住院天数和检查安排，不再仅仅按病情需要来定，也开始绕着费用达标来转。

2.2 第三个月：“编码的钥匙”被发现

真正的突破，来自一次意外的学习。科室编码员在和兄弟医院的编码员交流时，发现对方医院收治同样的重症肺炎，但平均结算率高出将近20%。差别不在治疗方案——两边的用药和住院天数差不多——差别在病案首页。对方首页上，“肺炎”被放在主要诊断，但同时伴有“低蛋白血症”“低钾血症”“轻度肾功能不全”等一串第二诊断。这些第二诊断，单独看都不算大问题，但合在一起，就能触发分组器把病例往上分一级——从“伴有一般并发症”升级为“伴有严重并发症与合并症”，支付标准相应提高。

编码员回到科室，在例会上小心翼翼地提了这件事。主任沉默了一会儿，说了一句极其关键的话：“既然人家能做，我们也能做。但是记住——不要造假。那些第二诊断，必须是化验单上确实有的。”

这句叮嘱在此刻是完全真诚的，在技术上也完全成立：那些低钾血症、低蛋白血症，确实在化验单上摊着。但一个从未被明说的位移发生了：一个常规补液就能纠正的轻度低钾，在过去从来不会出现在病案首页的主要诊断栏或并发症栏里——它太不重要了。现在，它的重要性被重新定义：不是被临床重要性定义的，而是被支付权重定义的。从这一天起，编码员开始频繁走进医生办公室，逐一检查出院病历的首页——“王医生，这个患者血钾3.2，你怎么没写上？”“张医生，呼吸衰竭是I型还是II型？写清楚，权重不一样。”

这仍然是“合理的”：因为病历本来就该完整。这个“本来”的边界在哪里，没有一个人说得清，也没有一个人需要说清——它只需要在每一轮查首页时被用眼神和语气传递下去。编码员在这个节点上承担了一个关键的“催化”角色：他不是一个只会录数据的文书，他是两个场域之间的“翻译器”，把财务逻辑的编码规则翻译成日常临床记录的书写指令。这个翻译器本身没有恶意，他只是在认真完成自己被要求做的工作——但这恰恰使得他成为发生链条中最隐蔽的齿轮：一句“写清楚，权重不一样”的叮嘱，在认知层面上把一个支付逻辑的指令，不动声色地塞进了临床书写规范的外壳里。

2.3 第六个月：操作在一轮轮反馈中走向精熟

到第六个月，围绕重症肺炎的一整套新操作已经基本成熟。住院天数被进一步压缩：入院前三天的急性期治疗基本保留；第四天起，一旦体温正常、炎症指标回落，即使患者自述还觉得“没力气”，也开始被安排“带药出院”，后续治疗转移到门诊输液。对于特别危重的、预估花费远超支付标准的患者，急诊科在收治时默许了一种委婉引导——“本院ICU床位紧张，建议转往省级医院”——这句话不违反任何成文的收治规范，但在统计上切实减少了那些“不划算”患者的入院数量。

而一个更精妙的变化发生在科室内部的教学环节。新入科的规培医生在第一次写病历时，被带教老师纠正：“你这个肺炎，有没有合并呼吸衰竭？血气分析看看，氧合指数多少？如果低于300，就可以下‘呼吸衰竭’的诊断。”“这个患者查了肝肾功能吗？如果肌酐有轻度异常，急性肾损伤也要写上去。”从医学教育的角度看，这是对病历完整性的严格要求。从支付逻辑的角度看，这是在教一个年轻医生如何“把该写的都写上”——而“该”字的边界，从这一刻起就已经与支付权重缝

合在一起了。这个年轻医生在几年后自己成了主治，带教下一批新人时，会说出完全同样的话。他从未在任何一个时刻被“教过”博弈——他只是在学怎么写好一份病历。教学的日常性，是这种内化最坚硬的护甲：它让认知置换看起来不像是被制度强加的，而像是从职业成长本身中自然长出来的。

2.4 一年以后：新常态定居

到这个科室实施DRG满一年，重症肺炎的平均结算率已经从不足70%上升到接近95%。月度报表上，亏损的大洞被补上了。但代价是悄无声息的：这一年里，重症肺炎患者的平均住院日缩短了近两天，出院后三天内非计划再入院率小幅上升（虽然从未被系统性地统计），科室收治的高危重症肺炎患者数量显著下降，这部分需要倾注更多心血的人被系统性地引流到了省级医院和被豁免了DRG考核的康复机构。

任何一个从外部走进这个科室的观察者，都不会看到任何“违规”行为。他会看到一群工作认真、经常加班、对患者态度良好的医生，在认认真真地写病历、查房、收治患者。他不会看见那只已经长出来的、与被它寄居的宿主外表无异的“新医生”——因为这只新医生，就住在每个老医生的身体里，而他们自己，已经感觉不到它的存在了。这不是任何一个人的错误。这是一个完整的发生过程：度量标准没有被“反抗”，也没有被“规避”——它被内化了，长成了一种新的常识。

三、“善意前提”的崩塌：为什么再精确的度量都会在落地时裂开

上面这间病房的故事，不是一个偶发的“执行走样”。本节退后一步，拆解故事背后的系统逻辑——为什么DRG的设计，不管多精细，都必然在落地时裂开一条度量永远照不到的缝。

3.1 制度制定者的“最坏情形”预想

所有支付改革的起点，都是一个明确的靶子：按项目付费时代，医院有多开检查、多用耗材、多住院天的内在冲动，因为每一张单子都是一笔收入。制度制定者对这个靶子的回应，可以被概括为三个步骤：第一，用一个打包价替代逐项计价，切断“多做即多收”的利益链；第二，用一套编码和分组工具，把模糊的“一次诊

疗”装进可度量、可比较、可考核的标准单元里；第三，建立一套监管和审计机制，查出“超标”和“违规”的行为并施以惩罚。

这套逻辑在各国DRG体系中以不同的技术形态反复出现。美国Medicare自1983年实施前瞻性支付制度（PPS）后，早期研究迅速记录了医院行为的系统性变化：住院日显著缩短，但部分缩减被转移到门诊和院后护理；编码升级现象在实施后的头五年内持续上升。欧洲各国在1990年代至2000年代陆续引入DRG后，也报告了类似的模式——分解住院、编码精度提升但伴有系统性的病情夸大倾向、重症患者被转往不受DRG约束的机构的概率上升。这些经验研究的共同结论是：费用确实控住了，但控费的方式远非制度制定者最初设想的“消除不必要的服务”——相当一部分“省下来”的费用，来自把成本转移到院外、转移到患者自付、或转移到其他未被DRG覆盖的支付类别。

在这些经验面前，制度制定者预想的“最坏情形”是什么？他们会告诉你：医院可能会分解住院——把一个病拆两次治，拿两次钱；可能会高套编码——把小病写成大病，拿更高的权重；可能会推诿重病人——因为重病人收费低、成本高。这些行为在他们看来，是制度的“漏洞”——就像篱笆上被兔子咬出的洞。他们的应对策略永远是同一个：补洞。分组不够细？拆得更细一点。权重不够准？用更精密的成本核算来校准。监管不够严？加大飞检频率、引入AI审核、建立黑名单。

这些应对策略共享同一个未经审查的前提：医院在做的，是“规避”制度——是在用制度自己的语言、在制度划定的游戏场上，寻找规则缝隙来获利。制度制定者预想中的博弈，是一群精于计算的理性行动者，在一套给定规则下，优化自己的策略以最大化收益。国际文献中主流的概念工具——从“供应商诱导需求”到“编码升级”“博弈均衡”——也都建基于同一个预设：行动者在算计，但他们算计的对象仍然是制度划定的那些指标。补洞本身虽然费劲，但它始终在制度制定者可理解、可操作的范围之内。洞被咬得再多，篱笆还是那道篱笆。

3.2 “善意前提”的实际内容

但这个看似合理的预设，其实偷偷塞进了一个远比“理性人假设”更深的命题。这个命题从未被明确说出，却构成了所有制度设计的默认地基。它可以这样表述：但凡行动者还在用制度划定的度量标准——费用、编码、权重、分值——来思考如何

做事，制度就仍然是那个规定游戏规则的“主人”。行动者也许会在主人的院子里乱跑，但不会把院墙拆掉。这就是本文所称的“善意前提”——不是指行动者道德上善良，而是指行动者默认留在制度划定的那根量尺的引力场内，用那根量尺的语言来回应制度。制度制定者与行动者之间，仍然共享一套可互译的“度量语言”。

这个前提之所以从未被质疑，是因为它在绝大多数治理场合确实成立。交通违法的司机在琢磨的是“怎么不被电子眼拍到”——他在用电子眼预设的“超速/不超速”这个二元分类来组织他的逃避行为。企业避税，也是在税法文本的条款缝隙里做文章——它在用税法自己的范畴（收入、成本、抵扣项）来重新描述自己的经营活动。在这些案例中，被治理者确实“留在制度的院子里”：他做的每一件事，都仍然可以被制度自己的那套分类体系完整识别、归类、判定。这正是“善意前提”被广泛默认有效的原因——它确实验证过，在太多领域都成立。

3.3 医疗的致命特殊性：医生不是兔子

但医疗与交通违法有一个根本差别：交警不需要懂内燃机原理，税法稽查员不需要会经营企业，他们只需要会读自己手里的量尺——速度读数、利润报表——就够了。而医生手里有一把自己的尺子，这把尺子不是医保局发的，是医学教育体系用五到十年的训练刻进他们身体里的。这把尺子的刻度不是费用、不是编码、不是权重，而是病史、体征、辅助检查、鉴别诊断、诊疗规范、预后判断。从阿伯特（Abbott, 1988）对专业管辖权竞争的系统分析，到莫尔（Mol, 2002）对医疗实践中不确定性处理的田野考察，都揭示了一个共同的事实：医生的专业判断不是一套可以被随意拆卸的附属技能，而是构成其职业认同的核心——它是这个行业在面对疾病的不确定性时，赖以做出集体承诺和个体决断的认知地基。换句话说，这是一种“内部尺子”——它不依赖外部权威的定期校准就可以自行运转。

当DRG把他们逼到“每一个临床决策同时也是一个成本决策”的墙角时，他们并不会停下来，在两个互斥的标准——“好诊疗”和“成本达标”——之间做痛苦的道德权衡。他们太忙了，每天要管十几张床，要应付病历质控、医保检查、科室考核，他们根本没有时间痛苦。他们用一套远比制度制定者预想更聪明的办法来应

对：不是用费用逻辑去“反”临床逻辑，而是用费用逻辑去重新定义临床逻辑本身——在脑子里把两套尺子的长法合在一起，长出一根新尺子。

这一过程不是“忙碌导致放弃反思”那么简单。组织社会学和认知心理学已经积累了丰富的机制来解释为何长期暴露于单一量化考核之下的专业行动者，倾向于从“痛苦权衡”滑向“认知缝合”。费斯廷格的认知失调理论早就指出，当个体反复做出与自己原有信念不一致的行为、且无法从外部找到充分的正当理由时，内心会产生强烈的不适；而消解不适成本最低的方式，不是改变行为，而是改变信念本身——让自己相信自己所做的事从一开始就是对的。班杜拉关于道德脱钩的研究则进一步描画了具体的认知操作：行动者通过委婉化重构（把“提前让患者出院”重述为“在现有条件下做到最好了”）、责任扩散（“整个系统都这么运作，不是我一个人能改变的”）、以及后果扭曲（“患者在家休养也许恢复得更快”）等机制，逐步将自己从道德不适中解脱出来。当这些个体层面的认知操作在同一个科室的同事之间被日常互动不断强化、验证、共享时，一种集体层面的认知重定义就发生了——用阿吉里斯与舍恩的术语来说，整个组织学会了围绕新的“使用理论”来运行，而那个被写在规章里、挂在嘴边的“信奉理论”，则变成了一种不必被当真检查的仪式性表述。

这就是“善意前提”在医疗必然失效的根本原因。兔子只能在篱笆上咬洞——它改不了篱笆。但一个经过十年训练、每天在生死一线做上百个判断的医生，他不是在一个在DRG的篱笆前绕来绕去找洞钻的兔子。他是一个被两套互斥的判断系统同时撕扯的专业行动者，而他的应对方式，不是在既有规则下优化策略，而是发展出一套新的、内部自洽的判断标准，让“对医学正确”和“对支付有利”在操作层面上变成同一件事。这一套新标准，制度制定者从来没有预想过它的存在，因为他们以为“医学正确”和“对支付有利”是两件不同的事——前者是专业规范，后者是经济激励——两件事可以冲突、可以权衡、可以妥协，但不会在认知的深层被缝合为一。

而一旦它们被缝合为一，制度就丧失了辨识它的能力。不是制度制定者能力不足，而是制度赖以辨识“违规”的那套分类体系——分解住院、高套编码、推诿重症——预设了“正确的临床实践”与“博弈化的临床实践”之间有一条清晰可见的边

界。而这套缝合手术恰恰磨掉了这条边界：一个被精心“完整化”的病案首页，既可以是一个负责任医生对患者复杂状态的精确记录，也可以是一个被支付逻辑整训过的编码员为拉高权重而做的叙事编织——两份病历可以长得一模一样，因为写它们的医生自己都分不清自己在做哪一件事了。当“违规”与“合规”在同一个认知动作中无法被分离时，制度不是被绕过去了，而是被彻底悬空了。

四、度量同构异质体：核心构造与发生机制

前面的病房叙事和逻辑拆解，共同指向一个被既有理论反复触及、却始终未能独立命名的现象。本节给出它的完整定义与内部构造，然后展开其发生机制，将“善意前提”的崩塌具体化为一系列可追溯的因果步骤。

4.1 定义与四个关键特征

度量同构异质体（metric-induced homomorphic heterobody）是指：当一套简化度量标准被嵌入一个已经拥有强大内部判断体系的专业共同体，并被确立为唯一的考核与资源分配依据时，该共同体的成员并非简单地“钻制度的空子”或“被制度扭曲”，而是会逐渐发展出一套与原有专业实践外表高度相似、内部认知质地却被彻底置换的操作惯习；这套惯习一旦形成，便不再需要高强度外部压力来维持——它被内化为一种新的职业常识，并通过日常的带教、考核与财务反馈在代际之间无声传递。[3]

这个定义有四个关键特征需要逐一说清。

第一，同构性。“同构”是指度量同构异质体在外表上与“正常的、合规的专业实践”几乎无法被辨别。病历上的每一个诊断都在化验单上有据可查；分解入院在入出院标准上都找得到医学上的合理依据；规避重症在形式上从未违反任何一条成文的收治规范。制度赖以识别“违规”的那套行政话语——分解住院、高套编码、推诿重症——预设了“正确的临床”与“博弈化的临床”之间有一条清晰可见的行为边界。而同构性恰恰踩在这条边界之上：它不造假，它只是从一堆真实的临床信息中，挑出对支付有利的条目放在聚光灯下，把无利可图的细节塞进阴影。一条边

界之所以能被称作边界，是因为两侧在某个特征上确实不同；而同构性所做的，是让两侧在这个特征上再也无法被分开。

第二，异质性。“异质”是指内部认知质地已经被彻底置换——具体而言，是指行动的主导目的与行动者的主观意图之间的边界已不可辨识。同一个行为——比如，一份“写得很完整”的病案首页——在度量同构异质体的内部，与传统规范化实践的內部，遵循着根本不同的优先序。后者首先是临床认知的真实沉淀，其次才是行政与法律的文件；前者首先是支付逻辑的合规文本，其次才是临床过程的记录。这种置换不是叠加——不是在一个不变的临床判断上再加一层支付考量——而是熔合：两套优先序被加热到无法分离的程度，形成一个新合金。写它的人自己，在日复一日的重复中，可能已经完全无法分辨这种倒置是从哪天开始的，因为在熔合完成之后，就不再有“两个可以分开的东西”可供比较。

第三，自维持性。度量同构异质体一旦形成，就不再需要高强度的外部压力来维持。它不依赖制度不断“拧紧螺丝”，因为它已经被内化为一种“我们本来就该这么做”的职业常识。一个新入职的规培医生被带教老师叮嘱“把该写的都写上，别漏了”——这是关于病历质量的善意提醒。但这个“该”字的边界是临床重要性还是支付权重，从未被明说，却在每一个月的科室盈亏报表中被精确地传递下去。当维持一个系统的成本从“持续施加外部压力”下降到“只需要让新手看看上个月的报表”，这个系统就已经不再依赖它的设计者而自行运转了。

第四，自扩散性。在区域竞争的压力下，任何一家不参与这套操作的医院都会在考核排名中处于劣势，从而被动地被拉进这套生态。但这种扩散不是制度主义所说的一般“压力下的趋同”。在经典的制度同构理论中，模仿常常指向合法性的获取。但在这里，扩散的直接驱动力是更硬的生存逻辑：当邻院的平均结算率高出整整十五个百分点时，这不仅仅意味着排名落后，更意味着科室奖金、人才吸引力、甚至下一年度与医保局的议价权全面处在下风。不跟进，就是被淘汰。合法性的趋同需要时间去沉淀，而生存压力的趋同只需要一张月度报表。

4.2 发生机制：从正常调整到认知置换的四步锁链

度量同构异质体不是被某一份文件设计出来的，也不是在某一时刻被某一个人“发明”的。它是在一个全部注意力被压缩进单一量尺的制度空间中，从无数次日常临床决策的缝隙里，自下而上地发生的。其发生机制可以用四步锁链来描述。

第一步：单一度量锁定全部注意力。当费用成为唯一可度量、可考核、可追责的变量时，科室的全部认知资源都被吸附到这一根量尺上。这不是说医生不再关心患者的健康——他们仍然关心——而是说，在制度分配注意力资源的机制上，费用是唯一被打上了高亮的项目，其他一切——并发症发生率、非计划再入院率、出院后功能恢复——都在注意力的暗区。认知心理学已经反复表明，被度量的东西会自动占据注意力优先序的顶端，而未度量的哪怕在价值上极其重要，也必定在决策中被系统性地边缘化。不是有人下令“不要关心长期功能恢复”——关心它的制度管道根本还没被铺设。

第二步：专业内部判断体系在暗区里开始自行寻找可操作的出口。医生的医学训练让他仍然能感知到那些被制度照不到的维度——一个高龄患者出院后独居的风险，一次大手术后在康复期缺乏家庭支持可能带来的功能衰退。然而在单一量尺的考核压力下，这些感知没有制度化的评估和记录渠道。它们只能在暗区里自行运作，寻找可以附着的对象。这里的关键微机制是一种“临床直觉的制度半衰期”：起初，医生仍然会凭直觉对这些暗区的信号做出反应，比如多留患者观察一天；但当这一天的住院日反复在次月报表上被呈现为“超支”时，这个直觉信号就逐渐失去了它在决策中的权重。它没有消失——医生仍然知道这个患者回家后可能会出事——但它不再能激活行动。直觉从“听到就出手”退化为“听到了，心里念一下，然后继续写首页”。

第三步：费用逻辑主动向这些被遗弃的直觉空间填入替代方案，认知合理化机制在此处发挥了决定性作用。这个过程不是直觉的消失导致行动空转，而是另一套逻辑填补了空出来的操作位置。当一个医生反复遇到“知道该多观察一天，但不能多观察一天”的矛盾时，他的认知系统会自己寻找缓解不适的解法——这一过程的微观机制，正是费斯廷格所描述的认知失调消解与班杜拉所分析的道德脱钩的联合运作。起初，医生每让一个他觉得还不够稳定的患者提前出院，内心都伴随一阵低度

的不适——这个患者回家后会不会出事？下次门诊见到他时情况会不会已经恶化了？但缓解这种不适最经济的方法，不是对抗整个支付体系——那需要极其高昂的个人代价且几无胜算——而是重新框定问题本身：不是“我被迫让一个还需要观察的患者出院”，而是“这个患者在现有条件下已经达到了出院标准，后续风险是患者自身病情演变所致，不属于本次住院的职责范围”。

这种“重新框定”不是一夜之间完成的，不是某一次晨会上的宣言。它是在一例又一例具体患者的处理中，被同事、被编码员、被月度报表反复暗示和强化后，一点一点地从“自我安慰”变成“真实信条”的。这个转变可以被理解为一个缓慢的自证预言：医生在外部压力下做出了一个临床上略有风险的决策，患者碰巧没有出大问题，他对自己说“你看，其实也没事”；下次他再做类似决策时，原本的那一丝犹豫就薄了一层；十轮以后，那层犹豫已经薄到他自己再也听不见。与此同时，科室内部围绕“如何智慧地处理DRG”的日常对话——晨会上分享的编码技巧、走廊里压低声音的经验传授、对新入科医生的“善意提醒”——构成了一个持续的、弥漫的合理化环境，使得每一个个体的认知失调消解都在集体的相互确认中被加固。至此，原本只是“应付一下”的策略性行为，就完成了向“这就是对的事”的认知熔合。

第四步：缝合后的新实践被考核制度“正向确认”，锁定为认知。这个被缝合后的新临床操作——在规范的外壳下完成了费用控制的实质——在次月的科室报表上会得到一个好看的数字：次均费用下降，结余率上升。制度给出的就是这个好看数字的正向反馈。正向反馈的力量不在于它奖励了行为——博弈论的强化学习模型也能解释这一点——而在于它允许做出这个行为的医生对自己讲一套“我没有在博弈”的叙事：我可以合理地论证这次提前出院是临床正当的，我的病历书写比过去更完整，我在编码上的严谨是病案质量的进步——而这一切都不是造假。正是在这个自治的叙事里，度量同构异质体完成了它的定居。从这一刻起，它不再需要外部压力来维持；它成了这个医生对自己职业角色的默认理解的一部分。

五、不可还原硬校验：它不能被已有的概念1:1替换

任何一个试图提出新概念的努力，都必须首先接受一个严苛的追问：这个概念是不是仅仅用新词包裹了一个旧义？它能否被现成概念无损失地替换？本节将“度量同构异质体”置于三个最接近的理论对手面前，不是为了证明其他理论“错了”，而是为了精确地厘定它切开了什么新东西。同时，为使每个不可还原论证都落到实证锚点上，本节为它们各提供了一个可观测的判别标准——一个可被后续经验研究操作化的代理变量。

5.1 与“制度性机会主义”的不可还原性

威廉姆森将机会主义定义为“以欺诈的方式追求自身利益”，该概念预设行动者内心保留了一个“真实自我”与“策略性表演”之间的清晰边界：行动者在表演，但他知道自己在表演——他知道自己是在“钻空子”。度量同构异质体与机会主义的根本差别，就在于这个边界的存毁。在度量同构异质体中，表演的自我与真实的自我不再能被切开：当一位带教老师对规培医生说“把该写的都写上”时，他并不觉得自己在教一个机会主义的技巧；他正在传授的，是他自己已经深信不疑的一整套“正确书写病历”的方式。

两者的不可还原性不只体现在理论叙述中，也反映在认知与行为的可观测差异上。一个持续处于机会主义状态的医生，面对编码争议时，其内部体验是“我知道我这么做是为了过关”——这句话本身就表明他仍能在认知上区分两套标准。而一个已形成同构异质体的医生，在面对同样的争议时，其典型反应是困惑甚至恼怒：“这个编码没问题啊，病情确实这么重”——他是真心认为，临床严重度与编码权重这两个原本独立的判断维度，本来就应该重合。因此，一个可操作的判别标准是：当被外部审计指出某条编码与临床记录的某些细节不完全一致时，机会主义者的第一反应是辩解策略的合理性（“这样写合规”），而同构异质体的第一反应是辩解临床事实的自明性（“这个患者确实有这个情况，你为什么觉得它不该写？”）。两者面对同一指控，调用的辩护资源来自两个不同的认知库——前者来自对规则的解读，后者来自对病情的认知。这个差异可以通过对争议处理过程的录音或访谈文本进行编码来观测。

5.2 与“扭曲激励下的博弈均衡”的不可还原性

博弈论可以很漂亮地解释，在单一费用度量、信息不对称与有限监管的条件下，医院与医保局之间会收敛到一个普遍采取编码升级、分解住院等策略的均衡状态。但度量同构异质体至少在三处超出了标准博弈论框架的射程。其一，博弈论将博弈规则视为外生给定，而度量同构异质体的运转会反过来改写规则——在DIP模式下，当年的“编码升级”拉高区域平均费用，推高下一年的支付分值，规则与策略共同演化。其二，博弈论均衡的“自实施”是一种策略均衡——每个参与者预期其他人维持策略，自己也就维持策略，但这种维持始终依赖对“别人会不会变”的持续计算；而度量同构异质体的“自维持”是一种认知均衡——一旦形成，就不再需要计算，因为“做这件事”和“做得正确”已经成为同义词。其三，博弈论不处理偏好内生化的深层过程：度量同构异质体不仅改变了医生在给定偏好下的策略选择，更改变了医生的偏好本身——它重新定义了“什么是一个值得认真对待的病例”。

这里的可观测判别标准，可以从决策的认知成本入手。博弈均衡下的策略性分解住院，需要医生在每一次决定“今天要不要让这个患者出院”时进行一次微弱的成本计算——想想DRG定额还剩多少、剩余住院天数够不够。这个计算可以是很快速的，但它仍然在消耗认知资源。而同构异质体形成后，这种计算会被一个简化的启发式完全替代——“这个患者稳定了，可以出院”。启发式替代计算，是被内化后的决策在认知外形上的显著特征。一个可操作的观测方法是：在给定的临床情景模拟中，评估医生做出出院决策的耗时和口头出声思维中是否出现对费用参数的提及。同构异质体决策更快、更少提及费用——不是因为医生有意识地“藏起来不说”，而是因为费用已经被内化在“临床稳定”这个判断本身的定义里，不再需要作为一个单独变量被调取。

5.3 与“惯习”的不可还原性

这是三个理论对手中最强的一个，也是不可还原论证需要花最大力气的地方。布迪厄的惯习理论几乎触及了度量同构异质体所描述现象的全部表层特征：一套感知、评判和行动的倾向系统，被行动者所在的场域所塑造，并通过实践被内化为第二自然。从表面上看，DRG支付场域中医生形成的“编码优化”“提前出院”等惯习，与医院管理场域中形成的“绩效考核达标”惯习并没有不同——它们都可以被

归入医生在当代医疗制度多重场域压力下形成的新惯习。那么，度量同构异质体凭什么不能只是一个被支付场域塑造出来的新惯习？

要厘清这个问题，不能从“场域逻辑是否统一”这个有争议的预设出发——布迪厄本人从未声称场域必须是逻辑统一的，文化生产场域中“纯粹艺术”与“商业艺术”的两极对抗恰恰证明场域内部可以容纳尖锐的分裂。更有效的切入点是布迪厄自己后期在《帕斯卡式的沉思》中对“分裂的惯习”的分析。布迪厄在此处讨论了行动者可能被不同场域塑造出多套倾向系统，并在特定场景下被同时激活、产生张力——一个出身工人阶级的学者在学术场域与原生阶层之间感受到的撕裂，就是典型的例子。这种“分裂”的核心特征在于：两套倾向虽然相互冲突，但始终作为可被区分、可被切换、可被反思的对象而存在。行动者能在不同场景下调用不同的惯习——在学术会议上引用理论，回到家乡的饭桌上切换语言——即便切换本身伴随着痛苦，切换的动作本身，恰恰证明了这两套惯习是可分的。

度量同构异质体所描述的现象，恰恰在这个可分的节点上与“分裂的惯习”分道扬镳。它不是在一个人身上安装了两个可以切换的操作系统，而是让其中一套操作系统（临床判断）的“输入端口”被另一套操作系统（支付逻辑）重写了。重写完成后，行动者不再能识别出“两套不同的标准”——他只有一套标准，而这套标准已经把支付逻辑融解在临床判断的内部，不再作为一个可独立指认的成分存在。这不是“两个惯习并存”，这是两个惯习的启动条件被同一条输入信号同时触发，长此以往之后，启动哪一个的判断本身就从行动者的意识中消失了。

这一区分的可观测判别标准，不能在日常操作中寻找——因为在日常操作中，同构异质体和分裂的惯习可能做出完全相同的外显行为。真正的差别，只能在两套惯习被强制分离的场景中暴露。具体的操作化方案是一个“角色剥离任务”：选择一个已实施DRG多年的科室，抽取若干被医保审计质疑过编码合理性的争议病例，在会议室内把当事医生重新放在这个旧病例面前，要求他完成两项独立的评估——“请暂时忘记DRG支付，只从纯粹临床角度，重新评估这个患者当时的情况以及你的处理方案是否完全合理。”随后再要求他完成第二项评估：“现在请加入DRG支付的考量，你的评估结论会有什么不同？”

在分裂惯习的情况下，医生可以完成这种剥离：他能够清楚地指出两项评估之间的差异——“如果只看临床，我会再多观察一天，但当时因为支付压力，我选择让他提前出院了。”两种叙事之间的差异本身，就是两套未能熔合的惯习各自独立存在的证据。在剥离过程中，他可能感到不适、痛苦甚至愤懑，但这些情绪恰恰证明了两套惯习的边界仍在——痛苦来自边界两侧的撕裂。而在同构异质体的情况下，剥离将无法完成或极其困难：医生会坚持他当时做出的选择无论在临床还是支付角度都是唯一正确的，他无法识别出那个“如果只看临床”的对比情景，因为他根本无法在认知上把临床标准和支付标准当作两件可以分别评估的事。他的反应不是“我能剥离但很痛苦”，而是“你这个任务本身没有意义——临床和支付本来就是一体的，分开看是外行人的想法”。这个剥离实验的困难程度——是轻松切换，是犹豫后还能分开，还是从认知根源上抗拒这个任务本身的合法性——构成了度量同构异质体与“分裂的惯习”之间的操作性判别尺度。后文第六节在回应“结果主义辩护”时，将会展示这个判别标准如何提供辨别两类侵蚀的线索。

六、沉默的侵蚀：回应两个最强反驳并显化其不可逆性

如果前文的诊断成立——度量同构异质体已在科室里长成了一种新常态——那么两个看似强硬的反驳立即浮出水面。第一个来自制度制定者最本能的自信：如果结果是合规的，那么过程中的认知重塑有什么问题？第二个来自医改争论中持续存在的尖锐批评：你这套“职业内核”“照护厚度”的说法，不就是医生群体维护自身利益和特权的修辞吗？本文必须正面回应这两个反驳，因为回避其中任何一个，都等于承认自己的核心论证站不住脚。

6.1 回应“结果主义辩护”：三种被忽略的不可逆侵蚀

医改政策圈内的“结果主义辩护”可以很精炼地表述为：“只要在既定规则下完成了有效的、合规的诊疗，医生内心的认知结构是否被支付逻辑改变，根本不重要。我们管的是行为，不是灵魂。”

这个反驳在自身的逻辑框架内是完满的。它不需要反驳本文前几章对发生机制的描述——它可以完全接受那些描述，然后说“所以呢”？要回应这个反驳，不能在结果主义的同一块地基上与它争夺“结果”的定义。必须退一步，追问结果主义框架

本身在评价改革后果时的默会边界：它只度量看得见的产出，而把所有看不见的损耗当作“不存在”。正是这个度量边界的设定，使度量同构异质体的三类深度侵蚀——照护厚度、职业内核、认知代际传递——全部落在评价体系的暗区中。这三类侵蚀的不可逆性各有不同，但共享同一个沉默特征：它们不触发报警、不被度量、不进入任何一张考核表格。

第一类侵蚀：照护厚度的不可逆剥离。本处的“照护厚度”指诊疗过程中超出最低必要标准、以患者长期功能恢复与社会回归为导向的那一部分照护投入——包括但不限于出院前的额外观察、对患者家庭支持能力的隐性评估、针对复杂并发症的跨专科协调、以及医患之间非正式的情感支持。这些投入在单一费用度量中表现为“多出来的成本”，因而在支付压力下首先被压缩。当一个重症肺炎患者提前出院，在报表上被记录为“病情好转，医嘱出院”时，一个正常运转的DRG考核体系不会追问：他在出院后的第一个周末发生了什么？他是否因为独居而无法在凌晨体温再次升高时及时就医？他是否因为“已经出院了就别再麻烦医生”的心态拖过了最佳的再干预窗口？这些追问不是道德追问，不是用“医生的良心”来加码；它们是实质追问——照护的厚度不在报表里，但在患者的预后里。

当一个科室系统性、长期性地在照护厚度上做出微小的、每一笔单看都似乎无伤大雅的减省时，这些减省会在整个医疗系统内形成单向的衰退：不是某一天突然崩塌，而是持续下探的基线被每一次月度报表的“正向确认”锁定为新的正常。这个基线一旦下移，要把它抬回去的代价远比当时把它压下来的代价大得多——因为抬回去需要同时顶住来自支付考核、科室文化和代际传承三个层面的压力。沉默的下降不可怕——可怕的是它不再被识别为下降。

第二类侵蚀：职业内核的静默改写。“我治的是病，不是首页。”——当一个在DRG落地之初仍能以这句话安放自己职业认同的医生，在三年后再也说不出这句话时，是什么东西被消耗掉了？被消耗掉的不是一种永恒的“职业本质”——本文无意暗示某个未被污染的黄金时代曾经存在。临床中“值得全力以赴的病例”这一判断力，是在特定历史条件下、在长期实践中由一代代医生与复杂疾病搏斗时发生的。它的发生条件包括：有足够的临床自主空间去接触和处理高危重症，有足够长的住院日去观察疾病演变的全貌，有足够稳定的团队去承担高风险决策的后果。当

这些条件被DRG持续压缩——最重的患者被转走、住院日被定额锁死、团队被财务报表的焦虑所笼罩——耗损掉的不是某种神秘的精神品质，而是这些条件所支撑的认知能力本身的再生产链条。丢失的不是一笔可以记账的损失，而是一种隐性的人力资本的断代。这个断代的后果不会出现在任何一份DRG考核报表里——它出现在十五年后，当这家医院的呼吸科再也找不到一个能独立处理高危重症肺炎的学科带头人时。报表和现实之间的时滞，是单一度量暴政最擅长的隐身术。

第三类侵蚀：认知的代际传递——这一笔不可逆性最强。前文第二-四节的叙事指向一个尤其值得担忧的后果——一个完整职业认知系统的自我复制和传播。一个在DRG环境中受训的医生，他所内化的所有临床规范和职业伦理，都已经是被度量标准重塑过的版本。当他自己成为带教老师，他会教给新一代医生的，自然也是这个已经被重塑过的版本。两代人以后，这个版本已经不再是一个“被重塑过的变体”——它就是临床规范的定义本身。而此时，任何试图恢复原初标准的努力，都会被认为是在推翻“常识”。这正是度量同构异质体最深的破坏力：它不但改变了当下的实践，而且通过代际传递将自己的“异质性”变成了下一代人的“正常性”。

6.2 回应“专业主义自我辩护”的反驳：内部判断体系的独立根基

上述对三类侵蚀的分析，必然会遭遇一个更尖锐的反驳。医改文献中一直存在一种批评：医生群体用“临床自主”“专业判断”的修辞来抵抗合理的成本控制。他们声称的“职业内核”“照护厚度”，不过是维护自身利益和特权的修辞——是用专业主义的美丽外衣，包裹着对旧日不受约束的自主权与经济利益的眷恋。从这个角度看，“度量同构异质体”不过是为这种利益辩护穿上了一套更精致的理论外衣。

这个反驳的力道在于，它不质疑本文描述的事实，而质疑本文的立场：你凭什么站在医生的职业认同这边，而不是站在控制医疗费用的公共利益这边？如果在两种叙事——一种是“医生在捍卫患者的照护厚度”，另一种是“医生在捍卫自己不受成本约束的特权”——之间，没有独立于双方立场的判别标准，那么本文的一切论证，都不过是对前一种叙事的一次精致的重述。

回应这个反驳，必须从医学教育的认知实质入手。有充分的理由认为，即便完全承认医生群体有维护自身利益和自主性的诉求，医生在长期训练中获得的内部判断体

系，仍不能简单还原为利益修辞。这种判断体系的核心，是一种不可代偿的认知能力——它不同于“随便谁都能权衡利弊后做出的偏好选择”。一个受过十年训练的重症医学科医生，在面对一个同时有呼吸衰竭和急性肾损伤的高龄患者时，他看到的不是一堆可以独立处理的症状列表，而是一个已经被他的训练整合为一种连续形态的“整体临床图像”。这个图像的结构——什么信号是关键、什么变化是危险的、什么组合指向某种特定的病理过程——不是他个人的“观点”，而是他的专业共同体在数代人的临床实践中沉淀下来的、被反复校准过的认知工具。

更重要的是，这个认知工具的维持，依赖于他持续接触足够多、足够复杂的病例。如果度量同构异质体导致的“系统性推诿重症”使得一个科室的医生逐年失去在复杂病例中磨炼这种认知能力的机会，那么损失的就不是医生个人的“成就感”或“职业特权”——损失的是这个认知工具在代际间的再生产。当十五年后的呼吸科主任面对一个真正的危重患者时，他做不出那个正确的判断，不是因为他不够努力，而是因为他的整个职业生涯中，从未被放在需要做出那个判断的场景里。这种损失是一笔真实的、可观察的公共损失——它伤害的是未来被送进这家医院的患者，而不是仅仅是医生群体的职业自豪。

因此，本文站在医生的“职业认同”这一侧，不是出于对医护群体的情感偏向，而是基于一个可被独立检验的命题：DRG所催生的认知置换，在系统性削弱一种对患者预后具有实质影响的认知能力，而这种能力的损耗无法被费用报表上的任何指标所捕获。如果有人想反驳这个命题，他需要论证的不是“医生群体也有利益诉求”，而是“重症患者被系统性推诿后，医生的复杂病例处理能力不会衰减，或即便衰减也不会影响患者的长期健康产出”。这个论证的举证责任在反驳方，不在本文。本文的立场不是“医生的利益高于费用控制”，而是“有一种费用控制的代价，目前未被度量——而这个代价，恰恰是费用控制本身在悄悄地推高的”。

七、出路：制造两种度量之间不可调和的持久矛盾

如果前面的诊断成立——度量同构异质体不是执行走样的偶然后果，而是所有依赖单一费用量尺的治理结构在高度专业化场域中的必然产物——那么主流的应对思路就需要被重新审视。

7.1 为什么更精细的精算不能破局

“DRG的分组还不够细，权重还不够准。只要把组拆得更细，每个组给更‘合理’的权重，就能收窄可操作空间，减少博弈。”这个主张是所有精细化管理者的第一应手。它在技术上是可行的：更细化的分组逻辑，更复杂的并发症与合并症调整，更多的权重因子，都在医保信息系统的算力支撑下可以逐步落地。但它有效界的边界极其狭窄。医学的复杂性不是“有没有并发症”，而在于同一个“低钾血症”，在这个患者身上是需要紧急纠正的严重电解质紊乱，在那个患者身上只是一个几乎不影响治疗路径的化验数值。分组再细，也不可能把每一个个体化的临床判断都编码进一个标准化的格子里。只要格子的边界存在，边界两侧就一定有灰色地带。

然而问题比这更深。度量同构异质体一旦形成，已经不再需要“钻边界”。它不需要去钻任何制度缝隙，因为在它形成的那一刻，“制度缝隙”和“临床正道”已经被熔合成同一个动作。更细的分组解决不了内化的问题，它只是在一条已经被置换掉的尺子上，划出更密的刻度。用更精密的算法去追赶一个已经把算法逻辑内化为主观判断的操作者，就像用一个更快的摄像头去拍一个已经学会了在摄像头前以假动作代替真动作的演员——拍得越清楚，假动作看起来越逼真。

7.2 引入第二种度量：制造矛盾，而不是回归本质

走出这个困局，需要做一件听起来简单、做起来极其困难的事：在费用这个度量之外，引入第二种度量——一种不以费用为尺度，而是以“患者在多大程度上因为这次诊疗而变得更好”为尺度的度量。以患者为中心的结局测量和患者报告的经历测量在英国、瑞典、荷兰的部分专科已有长达十余年的系统应用——一个髋关节置换患者在术后六个月内定期填写的疼痛、日常活动、生活质量标准化问卷，构成了该医院该病种“实际健康产出”的纵向记录。

但在这里，本文必须做一个根本性的澄清——这一点对于守住所提出的“度量同构异质体”概念至关重要。“引入以患者长期健康增益为定向的价值度量”，绝不能滑入一种“回归医学本质”的叙事。价值度量本身同样是一种社会技术物——它不是从永恒的理念天空降落下来的，而是在特定的制度条件、知识生产与权力博弈中被谨慎建构出来的。量表选哪几个维度、各维度如何加权、数据怎么采集和公开、

是否与支付挂钩——每一个环节都是一次价值选择，都嵌入了特定主体的利益与认知框架，都可能产生它们自己的意外后果与扭曲效应。单一价值度量长期独大，完全可能催生它自己的同构异质体——医院会学会如何“教患者填更好的分”，而不是如何治更好的病。已经有初步的国际证据表明，当患者满意度调查被纳入医院支付与排名时，部分机构会出现选择性收治（回避那些预期满意度较低但临床需求最迫切的患者），以及在有形服务体验（病房环境、餐食、等待时间）上过度投入、同时缩减无形但关键的临床照护资源的倾向——这恰好是另一个版本的度量同构异质体，只不过这一次被内化的是“服务体验”而非“费用效率”。

因此，真正要做的事，不是从一把“歪的尺子”换到一把“直的尺子”。真正要做的事，是让两把尺子同时存在，并且刻意让它们保持不可通约的矛盾关系。当前困局的根源，从来不是DRG的尺子不够准，而是它只有这一把尺子——所有的行动者都被迫在这把尺子面前跳舞，没有人能问“尺子照不到的暗区出了什么事”。引入一把独立于费用的、不以省了多少而以致好了多少为刻度的尺子，它的首要功能不是“纠正”费用度量，而是让费用度量失去垄断全部注意力的地位。

7.3 矛盾只能持续，不能被“解决”

当每一个科室同时面对两把尺子——一把追问“你省了多少钱”，另一把追问“你的患者一年后真的好起来了吗”——它就会陷入一种建设性的两难：不能光省钱，因为省钱省到影响了长期健康产出，价值度量会暴露它；也不能光花钱，因为花钱花出了超额的短期产出，费用度量会追究它的效率。正是在这种张力中，一种被两把尺度量过、被两把尺子共同逼出来的实践才有可能从内部生长出来——不是靠道德呼吁，不是靠更严的监管，而是靠两种度量之间不可被轻易调和的持久矛盾。

但这不等于说两种度量并存的制度就一定优于单一度量。必须诚实地面对一个风险：两种度量的并存，会不会仅仅催生一种更复杂的、能同时应付两把尺子的“双轨同构异质体”？这种担心不是理论上的杞人忧天。如果两把尺子的数据都掌握在同一批人手里，而他们又面临必须同时在两把尺子上“表现好”的考核压力，那么完全可能出现一种更精致的认知缝合——医生学会如何在压缩费用的同时，也学会如何“管理”患者的长期问卷：在出院时不厌其烦地叮嘱患者“下周会有人给您打电话问恢复得怎么样，您尽量选‘好’或‘很好’”，在随访节点前主

动联系那些可能有负面评价的患者进行“补救服务”。这种双轨异质体的危险性在于，它比单轨异质体更难被察觉——每种度量单独看都得到了正面反馈，而两套操作之间的缝合处，比单套操作的内部缝合更隐蔽。

这个风险的切实存在，反过来加强了本文的核心论证：出路不在于找到一把完美无缺的尺子，而在于永远不让任何一把尺子获得独存的地位。两种度量并存确实可能催生双轨同构异质体，但这不等于应该退回到单一度量——因为单一度量必然催生同构异质体，这一点前文五步已经给出了完整的发生学链条。关键的问题不是“两种度量是否完美无风险”，而是“两种度量并存的制度空间，是否比单一度量更有可能暴露、而非掩藏同构异质体”。本文的判断是肯定的——原因在于一个简单的机制差异：在单一度量下，被内化的博弈根本没有被“看见”的制度管道。而在双重度量且两者不可通约的格局下，当一家医院的费用指标排名第九但健康产出排名倒数第三时，这种“尺子打架”的局面本身就是一种信号。它不一定能自动瓦解已经形成的同构异质体，但它至少做了一件单一度量绝对做不到的事：它让两种判断体系的冲突变得可以被公开讨论，而不是被预先封在一个“合规”的盒子里。冲突的可见性本身，就是打破内化的第一步。在此之后的关键制度设计，不是去解决这个冲突——用一个“综合权重”把两把尺子合并为一把，那恰恰是重蹈覆辙——而是建立处理冲突的程序正义：当两把尺子对一个病例做出截然相反的判决时，由谁以何种程序来做权衡？这个程序的透明性、多方参与性和可被申诉性，比任何预先设定的权重公式都更重要。

这当然是一条漫长的路。从短期看，任何已经在现行博弈系统中获得稳定生存位置的机构，都没有动力去主动引入一把会打破自己优势的尺子。这把尺子的引入，更大的推力不大可能来自于体系内已经获益的机构，而更可能来自两头。一头是支付方本身——当医保基金的增长空间被进一步压缩，而费用度量的污染程度已达到“报表好看但实际控不住费”的程度时，支付方会有越来越强的动机去寻找一个独立于费用之外的、可以标定“真正健康产出”的度量工具。另一头是公众——当患者选择医院的依据从“这家医院便宜”升级为“这家医院治好的患者长期功能恢复更好”时，那些在价值度量上表现优秀的机构，就会获得一种比任何广告都坚固的口碑。这两头的推力——一个是财政压力，一个是信息透明化后公众的选医权——在历史上往往是推动制度变迁的两股最坚硬的力量。

制造矛盾，而不是发布标准——这是本文与所有“回归本质”论者最深的分野。不追求矛盾的解决，而追求矛盾的长期保持和制度化处理——这种立场本身，就是对单一度量思维的最彻底拒绝。不是去寻找一把好到足以一统天下的尺子，而是建立一个尺子之间互相质问的生态。

八、结语：打开一个比DRG更大的问题空间

本文的核心判断，可以凝缩为一句话：在一个只用一把简化量尺来度量高度复杂专业实践的治理结构中，度量同构异质体不是执行走样的偶然后果，而是系统中理性行动者在被剥夺了使用自身内部判断体系的空间之后，必然收敛的认知置换事件。走出这个困局，不能指望一把更精细的量尺，而必须引入一种独立的、不可通约于既有量尺的第二种度量，并刻意制造二者之间无法被轻易调和的持久矛盾——经两种度量共同校准的实践与制度安排，只能从这种矛盾中被逼出来，而不可能由任何预先设计的完美方案平稳推出。

这个判断不试图成为可以被标准实证主义框架“检验”的假说，但它确实提供了一个明确的可观察证伪条件：凡是在未来的医疗体系改革中，持续沿着“更精细的分组、更智能的审核、更严密的飞检”的单一路径深化、却迟迟不着手建构独立的、公开透明的长期健康产出度量体系的改革方案，其终局将不是消灭度量同构异质体，而只是将它驱赶到更隐蔽、更难辨识的认知缝罅之中。反过来，如果某一天，我们观察到有地区开始系统性试点“公开报告医院的长期健康产出数据并使其具有实质考核权重”，并且在这些试点中冒出了医院之间围绕“谁治得更好”而非“谁更会编码”展开竞争的苗头，那么本文的悲观诊断才有可能被现实超越。进一步的证伪是：如果在引入独立的健康价值度量并赋予实质考核权重之后，反向博弈系统不仅未消退，反而围绕着价值度量又催生出了规模化的“患者满意度填分辅导”或“选择性收治高满意度患者”等新型同构异质体，而医疗体系又无法通过持续引入新的、更不可操纵的度量维度来抑制这种内化，则本文关于“制造矛盾可以瓦解度量同构异质体”的核心论证，应被视为在实践中无法成立。

但这个判断更大的意图，或许不在政策建议的层面。本文真正试图打开的，是一个比DRG广阔得多的问题空间：在一个日益依赖简化量尺来做治理的时代——学术

论文的查重率、大学的就业率、城市的GDP增长率、社交媒体的点赞数——我们如何防止那些量尺，不论其最初的意图多么良善，最终反过来驯化了被度量者的人格与实践，直至度量本身所要守护的那个初衷被彻底掏空？度量同构异质体的发生机制绝不仅限于医疗支付场域。凡是在一个持有强大内部判断体系的专业共同体之上，被强加一套单一简化度量标准，而那套内部判断体系又被剥夺了表达与反制制度管道的地方，度量同构异质体都具备生长的土壤。当我们用“重复率”来度量学术诚实、用“就业率”来度量大学教育质量、用“点赞数”来度量创作价值时，我们都在各自版本的DRG中，量出各自版本的同构异质体——学者学会写“能过查重的论文”，大学学会“把灵活就业算进就业率”，创作者学会“创作更易传播却更空洞的内容”。而抵抗这种遗忘的唯一方式，或许不是去找到一根“更准的尺子”，而是永远不让任何一根尺子独存——让两把尺子互相质问，让笼子永远无法完全闭合。

附论一·最近邻切割（续）

本节由编辑部在发表前补写。用意只有一个：把原稿已切过的近邻之外、那些最可能整个吞掉本文核心概念的理论对手请上场，逐一走完「承认占位 → 剩余分工 → 方向相反的可裁决预测」三步。凡切不动的，如实写明切不动。

1. 与坎贝尔定律的分离：本文说的不是指标被腐蚀，是被度量者被替换成一个同形的新品种

承认占位。本文所述现象最强的现成占位是坎贝尔定律：任何量化社会指标越是被用于社会决策，就越容易受到腐蚀压力，也越容易扭曲和腐蚀它本欲监测的社会过程。次均费用与住院日在纸面上漂亮下滑而实质恶化，几乎是这条定律的教科书式演示。承认这一占位是必要的——若本文只是说「指标失真了」，坎贝尔早在1979年就说完了。

剩余分工。剩余的分工在于失真的落点。坎贝尔定律的因变量是指标本身的效度：数字与它所指代的现实之间的对应关系被拉断，其经典表现是造假、选择性上报、只教考点。它的隐含图景是一个仍然完好的实践被一层失真的数字所遮蔽——把数

字掀开，底下的东西还在。本文所论证的同构异质，其特点恰恰是数字没有造假：费用真的降了，住院日真的短了，指标与它所测量的那个量之间的对应关系完好无损。被替换掉的是生产这些数字的那个人——他的判断结构、注意力的分配方式、面对不确定时的第一反应，已经从临床推理换成了一种以定额为原点的推理。外形同构，内部质地置换。坎贝尔定律的处方（防造假、增加指标、交叉验证）在这里全部落空，因为没有造假可查。

可裁决预测。方向相反的可裁决预测：对实施 DRG 五年以上的科室，同时测量指标失真度（以病案首页编码与病历实录的不一致率为代理）与临床推理形态的位移（以病例讨论中「先想定额还是先想病人」的起始点分布为代理）。坎贝尔定律预测：两者应同向变动——腐蚀压力越大，造假越多、实践越走样。本文预测：两者可以分离，且最典型的科室恰恰是编码不一致率低（合规做得好）而推理起始点位移大。若不一致率低的科室其推理形态必然完好，则同构异质体不构成独立于坎贝尔定律的现象。

2. 与「善意前提」在委托代理理论中的正主的分离：本文否认的不是代理人的善意，是尺子的信息容量

承认占位。本文引入的「善意前提」，其正主是经济学的委托—代理理论：委托人无法直接观察代理人的努力，只能依赖可观测的产出指标，而代理人的目标函数与委托人不一致，于是产生道德风险。整套 DRG 设计正是这一理论的直接应用——用打包定额把医生的目标函数与控费目标对齐。本文所述的困境，很容易被归并为「激励相容约束设计得不够好」。

剩余分工。剩余的分工在于失败的定位。委托代理框架把失败归因于目标函数的错位与信息不对称，其解法是设计更精巧的契约——加入质量指标、引入声誉机制、延长考核期。它的隐含假定是：只要指标体系足够丰富，代理人的行为就能被引导回委托人所要的方向。本文要论证的是一个该框架无法容纳的命题：在医疗这样的场域中，「委托人真正想要的东西」有相当一部分在原理上无法被写进任何可核验的指标——一次判断是不是为这个病人做的，其证据只存在于做判断的那一刻的内部状态。这不是契约设计得不够精巧，是待度量对象的信息容量超出了任何可执行

契约的承载上限。因此本文不主张改良指标，而主张为不可度量的部分另设承认通道。

可裁决预测。方向相反的可裁决预测：追踪那些在 DRG 基础上叠加了大量质量指标（再入院率、并发症率、患者满意度、临床路径符合率）的医疗系统。委托代理理论预测：指标维度越丰富，激励相容程度越高，同构异质现象应当减弱。本文预测：指标维度增加会减少某些具体的规避行为（这一点委托代理是对的），但不减弱推理起始点的位移，甚至加剧——因为医生需要同时优化的约束更多了。若指标丰富度与推理形态位移呈稳定负相关，则本文的核心判断被推翻。

3. 与「制度性机会主义」的分离：本文的主角不是钻空子的人，是合规的人

承认占位。组织社会学关于制度性机会主义、脱耦与仪式性合规的整套文献，提供了另一条现成的解释路径：行动者在外部压力下采纳形式而规避实质，其表现是选择性执行、象征性合规与创造性遵从。DRG 之下的高编码、分解住院、推诿重症，几乎全部可以归入这一族。

剩余分工。剩余的分工在于：机会主义解释预设了一个仍然知道「什么是对的」并有意规避它的行动者——他心里有两本账，一本真的一本给上面看。这类行动者是可被治理的：加强稽核、提高违规成本、完善举报通道。本文所要指认的那个更棘手的形态里，行动者只剩一本账：他并非在临床判断与编码判断之间做取舍，而是那两种判断已经熔合为一种单一的、连他自己也无法拆开的临床—经济复合推理。他不觉得自己在规避什么，他觉得自己在好好看病。这正是「同构异质」中「异质」二字的分量所在——不是两本账，是一本被换过内容的账。稽核对此完全无效，因为没有可查的偏差。

可裁决预测。方向相反的可裁决预测：对同一批医生施以两类测试，甲为合规稽核（查编码与病历一致性、查分解住院），乙为盲法临床情境判断（给出与定额无关的匿名病例，测其推理起始点与不确定性容忍度）。机会主义理论预测：稽核不合格者应在盲法测试中同样表现出更强的经济导向，二者正相关。本文预测：两组显著分离——存在一批稽核完全清白、而在盲法测试中推理起始点位移最大的医生，且这批人往往被评为科室的模范。若这一分离不出现，则同构异质体可被机会主义框架覆盖。

附论二·证伪条件

以下条款由编辑部与原稿的自陈证伪条件合并整理。任何一条被严格的经验研究证实，本文的相应主张即应被修正或撤回。

1. 若在控制医院层级、专科类型与医生年资之后，DRG 实施年限不能预测临床推理起始点的位移，则同构异质体不构成一个由制度驱动的现象。
2. 若「推理起始点位移」这一构念无法被操作化为两名独立评估者可稳定编码的指标（例如基于盲法病例情境的出声思考编码），则本文核心停留在隐喻层。
3. 若合规稽核清白程度与盲法测试中的经济导向之间呈稳定正相关、从不出现本文所述的分离，则本文与制度性机会主义的分界线不成立。
4. 若在长期跟踪中，同构异质度高的科室其临床结局指标（再入院、并发症、生存）持续不劣于同构异质度低的科室，则本文所指认的「异质」可能并不构成实质损失，其规范性主张应被重估。

注释

[1] “善意前提”（good-faith premise）是本文提出的一个工作概念，用以刻画制度设计中被广泛默认的认知预设：行动者即便采取策略性应对，仍然以制度划定的度量语言为博弈框架。这一概念与“理性人假设”不同——后者仅假定行动者最大化个人效用，而“善意前提”进一步假定这种最大化过程不会溢出制度所设定的范畴体系。本文无意将这一前提普遍化到所有治理领域，而是聚焦于医疗支付这一高度专业化场域中该前提的失效条件。

[2] 材料说明：本文第二节所呈现的重症肺炎病房叙事，并非田野民族志意义上的实证材料，也不来自对某一家特定医院的实地观察。它是综合了中国多个DRG试点城市的公开政策文件、新闻报道、行业论坛中一线从业者公开发言等碎片信息后，重新整合而成的类型化叙事。该叙事的功能是作为“度量同构异质体发生机制”的思想实验性演示，不构成对任何特定机构或个人的事实描述。

[3] “度量同构异质体”造词来自“同构”（homomorphism）与“异质”（heterogeneity）的拼合。在数学中，同构指两个结构的形状相同但内部元素可以不同；在本文的语境中，“同构”强调被度量塑形后的实践在外表上与原有实践无法区分，“异质”特指行动的主导目的与行动者的主观意图之间的边界已不可辨识——行动者不再能够区分自己是在遵从专业规范还是在回应支付压力，因为两套优先序已在认知中被熔合为一个不可拆解的新合金。

参考文献

Abbott, A. (1988). *The System of Professions: An Essay on the Division of Expert Labor*. University of Chicago Press.

Bourdieu, P. (1980). *Le Sens pratique*. Minuit. [英文版: *The Logic of Practice*, trans. Richard Nice, Stanford University Press, 1990.]

Bourdieu, P. (2000). *Pascalian Meditations*. Stanford University Press.

Mol, A. (2002). *The Body Multiple: Ontology in Medical Practice*. Duke University Press.

Williamson, O. E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press.

——本次发表前补入（对应附论一各节，均为真实文献）：

Campbell, Donald T. “Assessing the Impact of Planned Social Change.” *Evaluation and Program Planning* 2, no. 1 (1979): 67–90.

Holmström, Bengt, and Paul Milgrom. “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design.” *Journal of Law, Economics, and Organization* 7 (1991): 24–52.

Bevan, Gwyn, and Christopher Hood. “What's Measured Is What Matters: Targets and Gaming in the English Public Health Care System.” *Public Administration* 84, no. 3 (2006): 517–38.

Espeland, Wendy Nelson, and Michael Sauder. “Rankings and Reactivity: How Public Measures Recreate Social Worlds.” *American Journal of Sociology* 113, no. 1 (2007): 1–40.

关于本文创新智商标注

本文的盲评分为 **135**（评分者未参与本文写作，具备评分资格）。附论一与附论二由编辑部补写，补写后的 **139** 是**修改设计目标**，不是认证分——按本栏规程，任何人不得为自己参与撰写的文本出具认证分。该分数**待独立复核**。评分口径为 SDE 创新智商五维（S 结构精确度／D 差异锐度／E 纠缠深度／I 不可还原性／F 可证伪性），参照语料与评分截至日随复核一并公布。