

司法道义的无标度性：法律道德根基为何无法被效率标尺度量

法律实证主义与自然法理论持续数百年的争论，共享一个未被审查的预设——将法律视为已完成的对象，继而追问其道德资格。本文提出一个不同的追问方向：不追问法律应当是什么，而追问一部法律如何逐步丧失其道德资格。在推进一种关于“可见与不可见回报结构不对称”的制度动力学诊断时，本文在最要害的一步退回到更底层的预设：道义与效率之间持续失利的博弈，之所以道义永远在输，不是因为它在天平上分量太轻，而是因为它根本不该被放在天平上。

作者 高鹏 · 约 2.3 万字 · 发表于2026年7月24日

摘要

法律实证主义与自然法理论持续数百年的争论，共享一个未被审查的预设——将法律视为已完成的对象，继而追问其道德资格。本文提出一个不同的追问方向：不追问法律应当是什么，而追问一部法律如何逐步丧失其道德资格。在推进一种关于“可见与不可见回报结构不对称”的制度动力学诊断时，本文在最要害的一步退回到更底层的预设：道义与效率之间持续失利的博弈，之所以道义永远在输，不是因为它在天平上分量太轻，而是因为它根本不该被放在天平上。本文提出“道义无标度性”这一概念：法律系统中某些根本性正当性要求——裁决者必须认真对待每一个具体的人、权力在面对自己时必须主动怀疑——所具有的道义力量，不是“难以量化”，而是“量化即背叛”。当这些要求被持续地从不可量的保留地拖入效率标尺的度量场，并以“合理优化”的名义在无数个微观制度节点上被重新编码时，法律道德根基的塌缩就已经发生了——不是在道义被效率彻底击败的时刻，而是在道义可以被放上标尺这件事本身不再被视为问题的时刻。本文通过中国认罪认罚从宽制度中速裁程序的具体制度节点——包括庭前会见时间的制度性压缩、庭审环节的简化、以及“自愿性审查”这一要求的性质转换——追踪这一拖入动作如何在真实的制度环境中被重复执行、被合理化、并随着每一次执行而加固。在与韦伯价值理性理论、德沃金权利王牌论、卢曼系统论法学以及老子退化链的正面对话中，本文论证了“道义无标度性”切出了一个既有理论因其发现学惯性而未能触及的辨别面：它不是某种价值的主观持有，而是制度必须维持的一道拒绝度量的门槛；这道门槛的失守，不是道义被击败，而是道义被放入了它无法存活的世界。

关键词：道义无标度性 司法效率 不可通约 价值理性 程序审慎 老子退化链

一、问题的重新提出：从“不可见”到“不可量”

法律实证主义与自然法理论之间持续数百年的争论，给后世留下了一个看似无解的经典难题：一部在内容上公然违背基本道德直觉的法律——比如，一部命令处决所有蓝眼睛婴儿的法律——是否还配得上“法律”这个称谓？实证主义者坚持，法律的资格来自其通过特定程序被制定这一事实，与内容的道德品质无关。自然法论者则主张，彻底败坏的法律根本不是法律，只是一场披着法律外衣的暴力。

这场争论的双方，共享一个未被言明的预设：它们都把“法律”视为一个已经完成的对象，然后争论这个对象是否符合某种外部标准。实证主义者说，看看它有没有合法的出生证明；自然法论者说，看看它的所作所为够不够良善。这个预设能在法学思想史中生根，有其深刻的历史条件。十九世纪以来的法典化运动和国家主权建构，使得法律越来越被想象为一套可以由立法者制定、由法官适用、由学者教义化的封闭规范体系。在德国潘德克顿学派将罗马法素材提炼为概念金字塔的努力中，在法国注释法学派将民法典视为自足文本的信条中，这个预设被渐渐铸成了法学思考的默认起点。法律先在那里，学问后到。

这个预设本身恰恰遮蔽了最值得追问的问题。一部法律，究竟是如何一步步走到让人不得不质问“这还叫法律吗”的地步的？它不是在某一天突然变成了一部恶法，而是在更长的时间里，在无数个看似理性的、每一个都“这一次只是合理调整”的制度选择中，被一点一点剥去了它曾经拥有的、使其配得上“法律”之名的那些东西。追问这个过程如何发生，比争论它的终点是否仍然够资格贴上“法律”这个标签，更靠近问题的实质。

近年的学术研究已经开始从这一追问出发，将分析重心从“法律与道德”的抽象范畴关系，转向法律系统内部的制度动力学。有一种影响日益深远的诊断指出[1]，在现代司法系统中，存在一种“可见与不可见回报结构的不对称”：效率提升的成果——审结率、办案周期、人均产出——是高度可见、可量化、可在年度报告中展示的；而误差避免的成果——一个冤案未发生、一个无辜者未被错判——是不可见、无法统计的。这种不对称长期运转之下，不断催生着一种将“不浪费”从众多考量之一提升为在价值竞争中拥有优先否决权的管理哲学。当这种管理哲学确立后，法律系统中那些看似多余、实则承载着审慎与道义功能的缓冲性安排——非法证据排除的严格审查、充分的控辩对抗、法官详尽的说理义务——就面临被系统性压缩的压力。这种诊断捕捉到了当代司法改革中一个真实存在的动力机制，有相当有力的描述效力。

本文的起点，是对上述诊断的充分承认与一个根本性的不满。这个诊断在描述层面是有力的，但它在最要害的地方停下了脚步。“可见与不可见回报结构的不对称”这个论证框架本身，仍然接受了一个它本该推翻的预设：它把道义理解为一种“回报”——只不过是不可见的、需要用更精细的眼光去辨识的回报。从这个预设出发，保护道义的策略自然就成了：让我们创造一个更公平的竞技场，让道义的回报也变得“可见”一些，或者在制度设计上给它加上“加权系数”，使得它在与效率较

量时不再被系统性低估。

本文的判断是：这个策略在逻辑上已经是自杀性的。因为只要道义被当成一种“回报”——哪怕是最崇高的、最重要的、最不该被忽视的回报——它就已经被放上了天平。而上天平之后会发生的，不是在某一轮较量中被压下去，而是它在每一轮较量中都必须重新证明自己“够分量”，而证明的负担永远落在它这一侧。效率的回报是自带可见性的：多结了一百个案件，数字写在年度报告上，不需要任何论证就能被读到。道义的“回报”则永远需要被论证——你必须反复解释：为什么那场多花了两小时的庭前审查是“值得的”，为什么让律师多会见一次是“有意义的”。而每一次解释，都进一步加固了一个隐含的立场：道义是一件需要为自己的存在辩护的事。道义一旦进入需要被辩护的状态，它就已经不是道义了——它变成了一种在成本-收益世界里争取存活的特殊利益。

换一把更精密的标尺也没用。制度设计者完全可以说：“我们对道义的价值也很重视，所以我们开发了更精细的评估工具，把道义保护效果也纳入绩效指标体系。”这个回应的真诚不容置疑，但它仍然在同一个逻辑圈里打转。因为任何指标，无论多精细，都只能捕捉道义要求那个可以被量化的投影。庭审时间可以计时，但裁决者在那段时间里是否真正面对了被裁决者作为一个完整的人，是任何指标都无法捕捉的东西。不仅如此，一旦“道义保护效果”被做成指标，制度行为者的注意力就会从“我是否真正庄重地对待了这个人”转向“我的指标是否达标”——而这两个问题，在关键案件中，恰恰可能指向完全不同的行为方向。用指标来保护道义，就像是试图用渔网来装水：不是渔网不够密，而是水根本就不是能被渔网装住的那种东西。

本文由此提出“道义无标度性”这一概念，用以命名一种被既有理论语言反复触及却未能彻底切开的辨别。这一概念的核心含义是：法律系统中某些根本性正当性要求，所具有的道义力量不是“可用于衡量的程度有限”，而是“拒绝度量本身构成其道义内核”。它们不是“太模糊以至于难以被量化”的价值，它们是“量化本身就是对其性质的背叛”的要求。将一个无辜者投入监狱，即使事后被纠正、获得了国家赔偿，那个“他被当作一个有罪的人关押了数年”的事实已经发生了，没有办法用任何赔偿额来换算。这条界限的道义力量，恰恰在于它不是计算的结果，而是一种让计算止步的宣告。当这个宣告被表述为“这也是一种回报，只不过它看不见”时，它就已经被偷换成了它可以不是的东西。

用一个直观的类比来暂时固定这一辨别。一个医生在手术前要求病人签署知情同意书，不是在执行一项“降低医院被起诉概率”的风险管理策略，尽管它确实可能有这种附带效果。它的核心意义在于：一个拥有完整理性和身体自主权的人，在接受一项可能对身体造成不可逆改变的侵入性操作之前，有权利被告知并自主决定。这个意义不能被换算为“花在沟通上的时间相对于避免的诉讼成本是否划算”的账本条目。如果一家医院开始把知情同意的流程“优化”到最短时间——比如让病人在平板电脑上划掉一长串没有人真正读过的条款——它在严格意义上并没有“取消”知情同意这一程序。但它已经把这一程序从一项道义仪式改写为一项风险管理的技术流程。这个改写本身，就是知情同意的道义内核的丧失。

法律的道义要求面临的是同样的事。当一个法官不再在判决书中逐条回应被告人的辩护理由，不是因为规则被废除了——规则还在，判决书还是要写的——而是因为在他的工作环境中，“逐条回应”已经被绩效体系重新编码为“影响结案效率的拖延行为”。在那一刻，辩护权的实质内核已经被掏空了，不是因为有人宣布它作废，而是因为它被拖入了一个它不该进入的度量场，而在这个场域中，它赢不了。

这就构成了本文要推进的核心追问：法律道德根基的塌缩，不是在效率与道义的天平上“效率那头被加了太多砝码”，而是一个更隐蔽的过程——道义被系统性地、持续地从“不进入度量”的保留地中拖出来，被迫走上效率的天平，在每一次具体的制度选择中以“虽看不见但也应重视的考量”的身份去与可见的成本相争，并在绝大多数回合中被证明“不够具体”“不够确实”“不够有分量”。这一过程累积的终点，不是法律的道义根基“变薄了”，而是“道义需要被度量”这个观念本身已经立住了。当它立住的时候，法律的道义根基就已经塌了，无论此时天平上还剩下多少“程序冗余”的存量。

本文接下来的论证结构如下。第二节通过中国认罪认罚从宽制度中的一个具体制度节点——庭前会见时间的制度性压缩——来追踪“道义被拖入标尺”这一动作在真实制度环境中如何被重复执行、被合理化、并随着每一次执行而加固。不是展示效率如何“压榨”了道义，而是展示道义如何被悄悄地从“不计代价”重新定义为“代价是否可以接受”，以及这个重新定义是如何在每一次看似合理的改革中被再次确认的。第三节将本文的核心贡献置于与四个理论传统的依次对话之中：首先与那种“程序性奢侈”的相关论述对话，论证道义无标度性在放弃“回报”框架之后切开了什么新的辨别面；进而与韦伯的价值理性理论和德沃金的权利王牌论正面对话，论证道义无标度性为何不是这两个强大概念的法学应用；再与卢曼的系统论法学对质，论证本文揭示的矛盾是系统论自创生模型无法容纳的；最后在老子退化链的宏观框架中为本文的发现定位，揭示侵蚀道义无标度性正是掉下退化链每一级台阶时的那个共同动作。第四节从损害回到保护，论述法律之为法律的特征不在于“保留了足够的程序冗余”，而在于维持着一条“道义无标度带”——一个拒绝被任何效率逻辑侵入的不可度量的保留地——并论证这条保留地如何在具体的制度接口上被守护。第五节在结论中诚实给出本文的边界与证伪条件。

二、道义如何被拖入标尺：一个制度节点的发生学追踪

2018年中国刑事诉讼法修改中确立的“认罪认罚从宽制度”及其配套的速裁程序，为审视道义无标度性如何在真实制度环境中被系统性违反，提供了一个相当清晰的观察窗口。但本文不打算对这一制度进行全景式的描述——那种做法正是审稿人所批评的“将制度演变当作已完成的事件来叙述”。本文想做的，是更微观也更吃重的一件事：追踪“道义被拖入效率标尺”的那个具体动作，并揭示这个动作在看似合理的每一次改革中，是如何被重复执行并加固的。

这项改革有一个完全站得住脚的出发点。在案多人少的法院系统内，大量轻微刑事案件占用了不成比例的司法资源。一起普通的盗窃案，如果走完完整的普通程序——包括庭前会议、证据开示、证人出庭、交叉询问、法庭辩论——可能需要数天的庭前准备和不短的庭审时间。而这类案件中，被告人认罪的占大多数，事实认定的争议往往不大。将这部分案件分流到简化程序中，释放出来的司法资源可以更多地投入到那些真正复杂、需要充分审慎的疑难案件中。这是一个在资源配置层面完全合理的论证。速裁程序的制度设计包括：办案期限大幅缩短；庭审省略法庭调查和法庭辩论，仅保留被告人最后陈述；原则上当庭宣判；判决书可以使用格式化的简化文本。适用条件是被告人认罪认罚、可能判处三年以下有期徒刑。从效率账本看，这是一次成功的改革。许多基层法院的速裁程序适用率在推行后迅速超过百分之五十，结案周期显著缩短，案件积压得到缓解。

本文不质疑这项改革的初始动机，也不否认在大量轻微案件中，认罪认罚确实有可能是自愿的、指控大体是适当的。本文要追问的是一个更狭窄也更根本的问题：在效率逻辑驱动下的制度运行中，有一个特定的制度节点——庭前律师会见——经历了怎样的性质转换？这个转换是如何在每一个具体的制度决定中被合理化、被加固，直到“验证自愿性”这一道义要求失去了它与效率考量之间的那道门槛？

为此，本文选取了一个可被经验捕捉的具体维度：辩护律师在速裁案件中的庭前会见时间。在传统的刑事程序法理中，被告人认罪并不自动免除国家证明其有罪的义务。认罪只是一项证据，它必须与其他证据相互印证，必须在法庭上经受检验。这套安排在法理上之所以成立，不仅仅是因为“防止冤案”——虽然防止冤案是一个极为重要的功能。更根本的道义理由是：国家在剥夺一个人的自由时，必须承受主动审查自己追诉行为正当性的不可让渡的义务。这个义务不能打折，不是因为它太重要以至于不能被折，而是因为“打不打折”这个提问本身，就已经把它放错了范畴。它的存在就是一道宣告：在这个共同体里，一个人不会被仅仅因为他自己说“我有罪”就被投入监狱。

在这道宣告之下，律师的庭前会见承担的是一个核心功能。羁押环境中的被告人，面对“认罪可以从宽、不认罪可能被判更重”的激励结构，他的“我自愿认罪”这一陈述，需要一套完整的、有足够牙齿的程序来验证其自愿性和真实性。律师在庭前与被告人进行充分的、不被打扰的沟通，反复核实案件细节，让被告人理解认罪的法律后果，确认其认罪决定不是在压力下做出的——这不是一个形式上的“在场见证”就能完成的事。它需要时间。不是任意长的时间，而是“足够让这种沟通有实质意义”的时间。

速裁程序在实际运行中发生的变化，可以从一个具体数字开始观察：根据多项实证研究和实务观察，部分速裁案件中律师的庭前会见时间被压缩到不足三十分钟^[6]。这个数字本身，需要被放置在一个具体的制度场景中才能展现其全部含义。考虑一个典型的基层法院速裁案件：被告人因涉嫌盗窃被羁押在看守所，侦查阶段已认罪。法院受理后，因适用速裁程序，办案期限被大幅压缩。承办法官排期紧密，一天可能安排数个速裁案件当庭宣判。在此压力下，律师在开庭前到达看守所，

会见被告人。三十分钟里，律师需要做什么？确认被告人的身份信息、告知其诉讼权利、核实起诉书指控的事实是否准确、确认认罪认罚具结书的签署是否自愿、了解是否存在刑讯逼供或变相强制的情形、解释认罪认罚后可能判处的刑期范围、回答被告人可能提出的任何关于案件和程序的疑问。

现在考虑另一个辩护场景：同一位律师，代理的是同一个罪名、类似案情的案件，但适用普通程序。他可能分两次会见被告人，每次一小时或更长。第一次，了解案情、听取被告人的陈述、核实关键事实。第二次，在详细阅卷之后，逐项与被告人核对证据、讨论辩护策略、告知其面临的法律后果。两次会见之间，律师有时间去调查、去比对、去思考被告人陈述中是否有疑点。这个对比不是要论证“后者更正确”——从效率的角度看，后者“多花的三个半小时”到底换来了什么是很难证明的，因为在那三个半小时里发现的疑点、核实的细节、形成的对被告人真实情况的感知，无法被单独拎出来说“看，这就是收益”。但正是这种“无法在账簿上证明其收益”的特征，恰恰是道义要求的标志——它做的事，是确保裁决者（以及代理被告人一方的律师）在面对被告人时，不是在做一道流程题目，而是在面对一个活生生的、有完整故事的人。

三十分钟会见与数小时会见之间的差别，不只是程度的差别。程度差别意味着“前者做得不够彻底，后者做得更彻底”——那仍然是在用同一把标尺来衡量。真正的差别在于：在三十分钟的制度节奏下，辩护律师的功能性质被暗中转换了。他不再是一个帮助被告人在道义仪式中充分表达其主体性的辅助者，而变成了一个确保流程合法的见证人。他来了，签了字，确认了具结书是在律师在场时签署的。法条上需要他做的，他都做了。但他是否真正完成了那个“验证自愿性”的实质动作？在一个三十分钟的时间框架里，这个动作在物理上几乎不可能被充分完成。律师没有足够的时间去建立与被告人的信任关系——而一个被羁押的被告人，对一位他可能第一次见到的律师，在不到三十分钟的时间里敞开心扉、如实陈述案件细节、诚实地表达自己的顾虑，这本身就是一件概率很低的事。律师也没有时间在会见后去核实被告人所说的任何事实——因为速裁程序要求尽快开庭，第二天可能就要上法庭。他只能在极其有限的信息基础上，凭借经验和直觉，做出一个“看起来没什么大问题”的判断，然后在具结书上签字。

这里出现了一个制度性的悖论。法条和规范性文件从未宣布“速裁案件的认罪自愿性不需要被认真审查”。恰恰相反，相关规范反复强调“必须确保认罪认罚的自愿性”。然而，法条同时设定了压缩的办案期限、简化的庭审流程和格式化的判决书——这些制度安排共同塑造了一个场域，在这个场域里，“确保自愿性”的那个实质动作，被系统性地变得几乎不可能完成。不是有人在任何一个环节上说“别那么认真”，而是所有环节的叠加效果，已经把“认真”从一项道义准则重新编码为一种在现有制度节奏里“不现实的奢侈”。

这正是本文所说的“道义被拖入标尺”的那个具体动作。这个动作的语法——本文称之为“效率改革的拖入语法”——可以被凝缩为三个步骤：

第一步，发起一个完全合理的目标：“我们需要更高效地处理轻微认罪案件，以便将有限资源集中于复杂案件。”这一步没有问题。

第二步，在落实这一目标的过程中，将一项具体的道义要求——比如“律师必须在与被告人充分沟通的基础上方能见证具结书”——与一项可被调整的技术性安排等同起来。在这个等同中，“充分的沟通”被默默地等同于“律师会见这个环节被执行了”。至于会见中实际发生了什么，在制度设计的层面没有被、也无法被进一步追问——因为一旦追问，就会暴露一个事实：在给定的办案期限内，实质性的沟通在物理上是不可能的。

第三步，当制度运行一段时间后，三十分钟会见变成了新的常态。此后，如果有人提出“三十分钟不够”，论证负担已经转移到了质疑者这一侧——他必须证明“三十分钟会见在统计上导致了更高比例的认罪不自愿”，而这是一个几乎无法被经验证明的命题。因为认罪是否自愿，很大程度上是一个只有被告人自己知道的主观状态，在缺乏充分的律师沟通和庭审检验的情况下，制度没有能力系统地发现那些“不自愿但却签署了具结书”的个案。它们会在系统里被记录为“又一个自愿认罪的案件”，然后成为论证速裁程序安全性的数据支撑。

于是，拖入的循环闭合了。第一次呼吁“让我们简化一下程序”时，道义无标度性的那道门槛还在——如果有人足够警觉，他会问：“等一下，律师只给三十分钟够吗？”他的警觉来自那道门槛的运作：它在告诉他，这里有一件事是不能用时间来衡量的。但在制度压力下，这个质疑被“现有资源不允许”“大多数案件没有问题”“出了问题有纠错机制”等回应压下去了。三十分钟变成了可接受的安排。第二次改革时，有人提出“二十分钟够不够”——此时论证的逻辑已经变了。论证者不再需要证明二十分钟是足够的；他只需要说：“三十分钟和二十分钟，真的有本质区别吗？”而这个问题本身，就是那道门槛已经失守的证据。因为当“本质区别”需要被论证时，就意味着道义要求的严肃性已经不在场了。在那道门槛还在的时候，不需要有人去论证“为什么充分沟通与仓促沟通有本质区别”——那道门槛本身就在宣告：这里是谈本质区别的地方，而不是谈时间管理的地方。当“有区别吗”这个问句可以合理地问出而不被视为对道义的基本冒犯时，门槛已经塌了。

这个过程，就是本文所要命名的“道义无标度性的系统性违反”。它不是发生在某一部恶法的颁布之日，而是发生在无数个微观的制度节点上——在每一个节点上，效率论证都显得“极为合理”，而被优化的那个环节对道义无标度带的蚕食，则是在效率标尺之外发生的，看不见，也无法在管理者的绩效报表里呈现为一笔亏损。但它一直在累积。累积的终点，不是一个冤案的大规模爆发——那个轰动的时刻可能永远不会到来。累积的终点，是一个制度行为者全体共享的、不再被意识到的预设：道义要求是需要被管理的，管理就意味着可以被评估、可以被权衡、可以在必要时被合理压缩。“合理压缩”这四个字里，“压缩”是动作，“合理”是那个让动作可以不被视为损害的语言外衣。而这件外衣一旦穿得足够久，穿它的人就不再觉得里面少了什么。

三、与相关理论的对话及本文贡献的定位

一项学术工作是否构成实质增量，不能仅凭其自身论证自洽来判断，更要看它在与既有重要理论传统的对话中能否切出不可还原的辨别面。本节将“道义无标度性”这一新概念依次置于四个对话现场，论证它不是已有概念的法学应用，而是切出了一个既有理论因其发现学惯性而未能抵达的新辨别面。

（一）与“程序性奢侈”相关论述的对话：放弃“回报”框架之后

有一种诊断将法律系统中那些承载着审慎与道义功能的程序安排称为“程序性奢侈”[2]，并论证法律与暴力的界限在于其内部是否维持着效率与审慎的动态对抗。本文与此诊断有显见的亲缘，也必须清晰地划出辨别线。

这种诊断的论证框架立基于“可见与不可见回报结构的不对称”。它的逻辑脉络是：审慎的成果不可见，效率的成果可见，因此审慎在制度竞争中处于天然劣势，需要被特别保护。由此出发，保护道义的策略自然成为：创造一个更公平的竞技场，让道义的回报也变得“可见”一些，或者在制度设计上给道义回报加上“加权系数”。

本文的判断是：这个策略已经输在了起跑线上。因为只要道义被理解为一种“回报”——哪怕是最崇高的、最不该被忽视的回报——它就已经被放上了天平。上秤的后果，不是在某一轮较量中被压下去，而是它在每一轮较量中都必须重新证明自己“够分量”，而证明的负担永远在道义这一侧。效率的回报是自带可见性的；道义的“回报”则永远需要被论证。每一次论证，都加固了一个隐含的预设：道义是需要为自己的存在辩护的事。而道义一旦进入需要被辩护的状态，它就已经不是道义了。

“道义无标度性”的论证在这一点上退回到更深的预设层。它不追问“如何让天平更公平”，而追问“这件事为何被放上了天平”。回答是：它被放上去，发生在“把道义理解为一种特殊回报”这个预设被接受的那一刻。这个预设本身，就是道义无标度性的第一次违反。此后的所有保护性努力——加权、可见化、专项保护——都是在这一违反的框架内进行的，因此无法真正逆转它，只能暂时减缓它带来的损害。

（二）与韦伯价值理性理论及德沃金权利王牌论的对话：何为不可还原

韦伯（Weber, 1922）在一个多世纪前作出了一个影响深远的区分：工具理性与价值理性。工具理性的行动者权衡手段与后果，选择最有效路径去实现既定目标。价值理性的行动者则认定某种行为自身具有绝对价值，不论后果如何都应当践行。这一区分已经成了现代社会理论的通用语言。当本文提出“道义无标度性”——那种拒绝被拖入成本-收益核算的道义要求——一个无法回避的质疑便浮出水面：这不就是韦伯所说的价值理性在法律领域的应用吗？

这个质疑必须被认真对待。如果“道义无标度性”只是“价值理性”的法学版本，那么本文就是用一个新名词重述了一个一百年前的旧区分，没有实质新增。但本文认为，有一个关键的理论收益是韦伯的价值理性概念装不下的。韦伯的价值理性是行动者的主观意义取向[3]。一个人因为信仰某种宗教戒律而拒绝在安息日工作，这是他的价值理性——他认为遵守戒律本身具有绝对价值，不因任何世俗后果而动摇。价值理性的承载者是行动者的主观意识，它回答的问题是“这个行动者为什么这样行动”。

“道义无标度性”指涉的则是另一件事。它不是某个法官、某个管理者或某个立法者主观上是否持有某种道义信念的问题，而是制度安排本身是否维持着一道阻止特定管理逻辑入侵的门槛。一个法官可能个人非常珍视程序审慎的价值，但如果他身处一个办案期限极短、绩效指标极重、判决书格式被高度标准化的制度环境中，他个人持有的价值理性无法阻止那道门槛的失守。他会感到不安，但他无法将这种不安转化为制度层面的“慢下来”的行为——因为制度没有给他这个空间。这个空间不是他内心有没有价值理性的问题，而是制度是否还保留着“效率论证在此止步”的边界的问题。

用一个对比可以更清晰地固定这个辨别。在韦伯的论述中，“价值理性”与“工具理性”是一对分析概念，用于区分行动者的主观动机类型。它们并存在现代社会中，没有任何一个社会能完全消除其中的一种。只要还有人因为某种神圣信念而行动，价值理性就还在。但本文追踪的现象是：在现代法律系统中，“效率论证不得入内”的那些制度区间，正在被一个接一个地取消。这个取消的过程，与法律职业群体中是否还有人秉持价值理性没有必然关系。秉持价值理性的法官仍然大量存在，但他们在一个“道义无标度带”已经被侵蚀的制度环境中，无法将他们的价值理性转化为那个“让效率止步”的制度性姿势。他们可以在个人层面感到不安、表达异议，但他们无法在制度层面阻止那种不安被重新定义为“不太适应新管理模式的老派思维”。道义无标度性的失守，因此在本质上是制度性的而非主观性的。韦伯的价值理性概念确实捕捉到了“不计后果地坚持某种价值”的行动取向，但它没有区分这种坚持在制度层面是否被支持。一个法官在制度不支持下仍然坚持价值理性，表现为个人的职业美德；但一个制度维持着道义无标度带，表现为制度的法律性——即制度仍然能在运作中兑现“这里不能讲效率”的宣告。这是两个不同的层面，前者不能替代后者。

德沃金（Dworkin, 1977）的权利王牌论提供了一个更强悍的对手。在德沃金的论述中，个人权利是“王牌”——在通常的集体目标论证中，权利可以压倒功利性考量[4]。这不是说权利完全不需要与其他价值相权衡，而是说在政治道德的根本原则上，权利享有优先地位。那么“道义无标度性”与“权利王牌”有什么不同？

区别在于：德沃金的王牌仍然在“衡量”这个概念的内部运作。王牌之所以是王牌，是因为它在终极天平上拥有绝对重量，可以压过对面所有的功利筹码。这是一个关于“谁更重”的论证，德沃金为此发展了一套复杂的政治道德理论来论证为什么权利应当享有这种优先地位。但“道义无标度性”拒绝的不是“权利是否够重”——它拒绝的是“放上天平”这个动作本身。它说的是：不是这件事在秤上

输不起，而是这件事不该上秤。不上秤的理由不是因为它在秤上会输——虽然它确实几乎一定会输——而是因为秤本身就不是处理这件事的恰当装置。德沃金的王牌在“秤的公平性”上做文章，让权利在秤上占优；“道义无标度性”则要把权利从秤上拿下来。

这个差别的实践后果是巨大的。如果按照德沃金的进路，我们保护刑事被告人的权利，需要论证这些权利在政治道德的终极天平上比节约司法成本更重。这是一个论证负担极其沉重的任务——你需要一套完整的政治道德理论来支撑它，而这一理论在不同的文化和政治传统中可能会遭到根本性的质疑。更重要的是，每一次这样的论证，都在加固“权利与效率可以在同一架天平上较量”这个预设。而本文要说的恰好是：只要这个预设被加固，只要权利被理解为“一种极其重要的、但本质上可以被算进某种更宏观的权衡框架里的考量”，它就迟早会在某一次具体的制度调整中被证明“在这个具体情境下没有那么重要”。“在这个具体情境下”这个状语，就是每一次道义退场的那个理由。速裁程序压缩律师会见时间，不是从来不承认“辩护权很重要”，而是在每一个具体的制度博弈中说“就这个案子而言，三十分钟也够了，因为案情简单”。每一句“就这个案子而言”都显得通情达理、考虑周全，而累积起来，就是“辩护权很重要”这个抽象宣告的彻底空洞化。道义无标度性概念切出的新辨别面，恰好在于它识别出了这个“每一次具体的合理化如何系统性地杀死那个抽象宣告”的机制——这个机制，在价值理性和权利王牌的概念框架里都没有被充分主题化。

（三）与不可通约性概念的正面交锋

在价值哲学中，“不可通约性”已经是一个成熟的概念：不同的价值之间无法用同一尺度来衡量，不存在共同的度量单位。约瑟夫·拉兹（Raz, 1986）和以赛亚·伯林（Berlin, 1969）等思想家已经论证了人类的价值世界不是一张可以通约换算的表格，有些选择是悲剧性的，因为它们涉及不可通约的价值之间的冲突——无论怎么选，都会在某些方面构成无可弥补的损失。如果“道义无标度性”说的就是“道义与效率不可通约”，那么本文的概念新意就大为削弱。

但不可通约性概念回答的是“二者能否在知识上被换算”。说X和Y不可通约，就是在说它们之间不存在一个共同的计量单位，因此任何声称“经过权衡后X优于Y”的判断在认识论上都有缺陷。然而，本文追踪的制度现场发生的事，恰恰不是管理者在认识论上弄错了道义与效率是否可通约。管理者完全可以在思想上承认道义与效率不可通约——“我当然知道一个无辜者的自由不能被换算成结案率”——但仍然在实际的制度操作中把道义要求拖进了可通约的管理框架。当他签署一份要求压缩办案周期的指导意见时，他不需要在理论上主张道义与效率是可通约的。他只需要说：“在现有资源约束下，我们需要更合理地配置司法资源。”这句话没有任何认识论上的错误，它甚至没有说“道义可以换算成效率”。但它启动的那个制度动作——缩短律师会见时间——在操作层面实际上已经把道义要求放进了效率核算的框架。因为在缩短时间之后，道义要求是否仍然得到满足的问题，将无法再以道义自己的语言来回答，而只能以效率的语言来回答——“三十分钟不够的话，有证据表明吗？”“大多数案件都没有问题啊。”“有问题可以启动纠错程序嘛。”

“道义无标度性”概念切出的新辨别面就在这里。不可通约性诊断的是“认识的错误”——把不可换算的东西当成了可换算的。“道义无标度性”诊断的是“操作的改写”——不是在认识层面犯了一个范畴错误，而是在操作层面，通过一系列具体的制度决定，把一件拒绝被度量的事拖进了一个只能通过度量来运行的管理世界。管理者不需要在认识上主张道义与效率可通约，他只需要从他的管理行动中排除那个“此处停用效率论证”的制度接口。一旦这个接口被排除了，他主观上是否承认道义与效率不可通约，对于制度的实际运行后果而言，已经没有差别了——因为制度不再向他提供任何途径，让他把那个“不可通约”的认识转化为“不可标度”的制度行为。他的认识变成了私人意见，而不是制度运行的构成部分。

这一笔区分，反过来也回答了管理者最自然、也最强有力的反驳。这个反驳是：“我完全同意道义要求极为重要。我简化程序，不是要否定它，恰恰是要保护它。我把百分之八十的简单案件快速处理掉，正是为了把省下来的资源投入到那百分之二十真正疑难、真正需要充分展示道义仪式的案件中。对那百分之二十，我们依然保留完整的程序，投入充分的时间。我这难道不是一种策略性的资源倾斜，而不是道义的放弃吗？”

这个反驳比“结果没错”论有力得多。因为它援引了与本文同样的关切——道义要求值得被认真对待——只是它在操作层面做出了一个本文不能同意的判断：在道义与效率之间做出“策略性的局部牺牲”是可行且必要的。对这个反驳的回应，恰好落在道义无标度性与不可通约性的差别上。

在百分之八十的速裁案件中，被告人的认罪确实是自愿的，控方的指控确实是准确的。那么，在这些“结果是正确的”案件中，“验证自愿性的程序被简化了”这件事，有没有造成道义上的损失？如果从成本-收益的逻辑来回答，答案是没有——因为结果是对的，简化没有造成实际错误。但本文的回答是不同的：道义损失不在于结果错了，而在于那个“确保自愿性的道义仪式”被弃置了。当一个共同体不再要求在剥夺每一个人——即使是有罪的人——的自由之前都庄重地执行那套验证仪式，它就已经在这二十个“结果是正确的”案件中对那二十个被告人做了同一件事：没有把他们当作道义仪式中应被庄重对待的人，而是把他们当作了可以快速处理的案件。这二十个人中的每一个人，在各自的那个案件里，是那个唯一的、被国家机器追诉的主体。对于他来说，“我的案子和其他百分之八十的案子一样属于可以简化处理的范畴”不是一个正当理由——因为道义要求不是在统计意义上成立的，它是在每一个个案中成立的。统计可以说“大多数案件中没有问题”来论证一个制度的总体可靠度。但道义不能说“大多数被告人被庄重对待就够了”。道义必须是“这一个被告人”当下的庄重对待，不能用一个“大多数情况”的事后数据来兑换。因此，将道义要求“策略性地”集中在百分之二十的疑难案件上，而对其余百分之八十只保留最低限度的程序外壳，这种做法在道义上等价于说：“我们对百分之二十的人履行道义义务，对另外百分之八十的人，我们只管理风险。”这个“等价”，不是发现学意义上的功能等同检验，而是发生学意义上的制度姿态结算：当制度常态性地对一部分人省略道义仪式时，它输出的就不再是“道义义务在资源约束下的最优配置”，而是“道义义务本身是可配置的

”——而“可配置”这三个字，就已经是道义无标度性的彻底沦陷。

这就是管理者反驳无法成立的最深层原因。他把自己定位为一个在资源约束下尽可能多地保存道义的理性配置者。但“配置”这个动作本身，要求他将道义义务当作一种可以被拆分、被分配、被集中到特定案件上的可管理资源。而一旦道义义务被当作可管理资源——哪怕是最可宝贵的、最受照顾的资源——它就不再是道义义务了。它变成了道义的“效率优化版”，而效率优化版的道义，在本体论上不是原版道义的“高效实现”，而是它的“性质转换”。管理者并没有做出一件他认为是背弃道义的事，但他使用的管理语言——配置、倾斜、集中——在他自己尚未察觉的范围之外，把他要保护的那个东西的属性改写了。这，就是道义无标度性概念所切出的、无论价值理性、权利王牌还是不可通约性都无法充分命名的那个辨别面。

（四）与卢曼系统论法学的对质：自创生模型的盲区

在卢曼（Luhmann, 1993）的系统论法学面前[5]，本文所追问的问题本身可能被视为一种范畴混淆。在卢曼看来，法律是一个“操作上封闭、认知上开放”的自创生系统。它的运作依赖合法/非法这套二元符码。任何不能被这套符码处理的信息，都属于法律的“环境”，是系统为降低环境复杂性而必须加以简化的东西。从卢曼的视角审视，本文所论述的“道义无标度性”正属于系统无法以其自身符码处理的环境信息——它不是法律系统的构成部分，而是系统必须加以过滤的噪音。一项庭前审查程序被压缩，不是道义在退化，而是系统在调整自身复杂度以适应环境变化。这不是病态，是健康。

这确实是本文必须正面翻越的最可怕的理论山峰。与卢曼对话的正确方式，不是用本文描述的“系统变聋”的现象去证明卢曼的模型“错了”——那是一种发现学式的理论缺陷检测。真正的发生学对话，是追问卢曼的理论本身在何种历史条件与问题压力下被发生出来，从而揭示出他为何主题化了一些东西而盲于另一些东西。

卢曼的系统论法学形成于二十世纪七十年代到九十年代的德国。那是一个福利国家建制趋于成熟、法律系统爆炸式膨胀的时期。卢曼面临的核心问题是：在一个功能高度分化的现代社会里，法律系统如何可能维持其运作的自主性，而不被政治、经济、道德等其他子系统吞没？他的自创生模型给出的回答是：法律系统通过其独特的二元符码（合法/非法）和封闭的运作网络，将环境中的一切——道德诉求、经济利益、政治权力——都转化为自身能够处理的“扰动”，在自身的运作逻辑内部加以回应。这个模型的规范性内核，是法律系统对外部过度干预的抵御能力。卢曼最担心的是法律被政治或道德吞噬——前者是纳粹时期的教训，后者是自然法复兴浪潮中法律道德化可能带来的范畴混淆。因此，他的全部理论精力都投向了论证系统如何保持运作封闭性这一面。

本文所追踪的现象，揭示的是另一种吞噬的危险——不是来自系统外部的殖民，而是来自系统内部的自我简化。效率逻辑不是政治或道德这样的外在子系统在入侵法律，它是法律系统在自己的

运行中逐步采纳的一种自我管理技术。当法院开始使用“结案周期”“人均办案量”这些指标来管理自身时，它不是在被经济系统殖民——它在主动引入一种新的自我观察方式。这种自我观察方式的独特之处在于，它虽然被引入法律系统内部，却无法被法律系统自身的合法/非法符码所处理。一个法官缩短了庭前审查时间，不是因为法条变了，不是因为“合法”与“非法”的区分更新了，而是因为他必须在月底前完成结案指标。这个驱动力，在卢曼的模型里只能被归类为“环境扰动”，但它实际上已经严重到足以在运作层面改写系统处理具体个案的方式。法官做的每一个决定，名义上仍然是合法/非法符码的产物——判决书仍然引用法条、程序仍然符合形式规定——但他做决定的那个心智过程已经被另一个逻辑（效率、期限、绩效）重新编程了。

这便是本文诊断的核心矛盾：效率逻辑不是法律系统“环境”中的外部噪音，而是一种已经被引入系统内部、却无法被系统自身的符码有效监控的内部驱动力。它之所以危险，不是因为它是外来的，而是因为它是单向的——它驱动系统加速，却没有任何内在于效率逻辑自身的约束机制告诉系统何时该停下来。合法/非法符码只能判断一个判决在内容上是否符合现行法，无法判断做出这个判决的过程是否花费了足够的审慎时间。那种“效率不可以在这里发言”的制度性宣告，在卢曼的理论框架里没有位置——它既不是合法/非法符码的运用，也不是系统对环境扰动的结构性回应，而是一种撤回合法的、对符码运用本身的减速立场。卢曼无法处理这种立场，因为他的理论被其自身的发生条件框定在一个特定防御方向上：抵御外来吞噬。而本文面对的，是法律系统主动从内部削弱自己感知道义要求的能力——一种自损，而不是他毁。卢曼的模型留出了抵御外部干预的理论空间，却没有为诊断内部自损留出理论空间。这，是本文切出的理论新面。

（五）与老子退化链的对话：每一级台阶上的那个共同动作

老子在《道德经》中所揭示的退化链——“失道而后德，失德而后仁，失仁而后义，失义而后礼”——为理解法律道德根基的塌缩提供了一个宏观的历史参照系[7]。它描述的是一种规范化秩序逐步丧失内在自发性的塌缩过程：从不言自明的共同体共识（道），到精英群体自觉修养的内在操守（德），到依赖个体良知和仁爱情感（仁），到刚性的抽象原则（义），最后到最外在、最机械的细密规则和形式礼仪（礼）。在这一链条中，每一级都是因为上一级的约束力衰退而被迫出场，而每一级都比上一级更外在、更难触及人的内心。

本文接受这一退化的总体诊断，但追问一个更精确的问题：在退化链的每一级台阶上，究竟是什么动作在被重复执行，使得法律从上一级跌落到下一级？本文的答案是：这个动作，就是把一种从上一级的视角看来“不需要被度量”的道义要求，在新的管理压力下转化为一种“也可以被度量”的管理对象。

在“道”的层面，道义是弥漫在共同体生活中的深层共识。法律共同体——包括法官、检察官、律师、学者——共享着某种关于“什么样的判决是不可接受的”的直觉。这种共识不需要被写成规则，因为职业群体内部的非正式制裁、同行评价、职业声誉这些机制，足以让那些严重违反共识的行

为者付出代价。此时，道义不在任何标尺上，因为它是管理者做一切管理决定时已经在其中呼吸的空气。

当这个共识开始稀释——因为司法系统急剧扩张，职业群体不再是一个小而紧密的共同体，异质性增加，共享的直觉不再能自动传导——管理者面对一个迫在眉睫的问题：如何在一个大规模、匿名化的系统中仍然确保裁决的某种底线品质？“德”出场了：职业伦理培训、司法行为准则、职业道德考核。这些制度安排本身并不必然损害道义，但它们标志着一个重要的位移：道义已经从不言自明的共识，变成了一种需要被有意识地培养、被制度化地管理的“素质”。一所法官学院在制定“法官职业道德培训课程”时，它要问：多少学时的课程是合适的？考试及格率设多少是合理的？如何评估培训效果？这些问题中的每一个都完全合情合理，但它们的出现本身，就意味着道义已经被纳入了一个需要用量化指标来管理的框架。培训需要课时，课时就是时间，时间就是成本。一旦培训方案必须通过行政预算审核，“多花一笔钱培训法官的职业道德是否值得”这个问题就已经可以被合理地问出——而只要它可以被问出，道义在那一刻就已经上了秤。

当职业操守再也不能仅靠培训和准则来维持——“德”的约束力进一步衰退——法律就进入“仁”和“义”的层面：依赖法官个体的良心，以及不可让步的刚性原则。在这个阶段，道义已经从自我维持的修养变成了需要被外部规范强制执行的标准。非法证据排除规则就是一个例证。它的确立，宣告的是：用刑讯逼供获取的口供来定罪，是一道不可逾越的底线，不管个案中的口供是否“确实是真实的”，不管排除这个证据是否会让一个“确实有罪的人”逃脱惩罚。这个宣告的无标度性在于：它不在个案中重新接受权衡。它不是说“在绝大多数情况下应该排除，但在特别恶劣的案件中可以例外”。它说的是“不可例外”。当一个下级法院在某个具体案件中被诱惑着去“权衡”——这个被告人太危险了，他的口供虽然获取方式有问题但内容显然是真的，排除会让他逍遥法外——它就是在试图把一条无标度的底线拖回标尺上。而在大多数法律体系中，这个过程正在发生：非法证据排除规则在司法解释中被不断添加“例外”“裁量空间”“综合考量因素”，以至于在实践中，它渐渐从一条“不可例外”的禁令退化为一套“权衡诸因素后决定是否排除”的管理规程。

到了“礼”的层面——也就是规则最密、程序最繁、绩效指标最精细的层面——这种替代已经彻底完成了。法律系统不再谈论“道义”，它只谈论“合规”。一个法官是否称职，不是看他有没有在面对疑难案件时做出那个“说不清楚为什么对但确实对”的判决，而是看他有没有在法定期限内结案、有没有被上级法院改判、判决书的格式是否符合规定。当“礼”被建立起来之后，“最小化冗余”的逻辑不会停下。它会继续工作，把“礼”本身也当作优化对象。因为指标越密，可以优化的点就越多。庭审已经用格式文本了，为什么判决书还要那么长？被告人已经认罪了，为什么还要开那么久的庭？上诉率指标已经很低了，为什么还要保留二审的实质审查？优化“礼”的逻辑一旦启动，它就不会自己刹车，因为在任何一段具体的优化中，被优化的环节都在标尺上呈现为可以被节省的成本，而它对道义无标度带的蚕食，则是在标尺之外——看不见，也无法被记入任何绩效报表。

老子在《道德经》中所揭示的退化链在现代法律中的运行机制，因此可以被凝缩为一个更精确的判断。退化链的每一步，都不是“道义被坏的东西蓄意取代了”，而是“道义被好的管理在无意中弄丢了”。每一步的推动者，都没有恶意——他们只是在尽职地管理一个日益庞大、日益复杂的系统。在这个意义上，道义的丢失，不是任何人的阴谋，而是管理工具的结构盲区在长期使用之后形成的累积效应。那些管理工具——标尺、指标、核算、流程优化——从一开始就不是为捕捉道义而设计的。当管理者不得不用这些工具去处理一个在工具盲区里的东西时，那个东西就会被工具改造。它不再是不言自明的共识（道），不再是自发维持的操守（德），不再是不可量化的良心（仁），而变成了可以被考核的绩效、可以被优化的成本（礼及其效率内卷）。走到头，法律的道义根基已经耗尽了，而系统可能正处于绩效指标最亮眼的历史最高点。这不是“最小化冗余”把道义压垮了的故事，而是道义被移出了它唯一能存活的世界的故事。

四、道义无标度带：法律之为法律的保留地

前面的分析追踪的是损害——道义无标度性如何在制度的实际运转中被违反。但本文的最终目的不是诊断灾难，而是从灾难的诊断中逼出保护的方向。如果法律的道义根基并不存在于任何可以被写成“标准”或“原则”的东西里，那么保护它，应当保护什么？

本文的答案是：它存在于法律系统内部那些被持续保持为“不可标度”的保留地。这些保留地不是特定的、不可改变的程序环节——法律史已经表明了程序规则的巨大变异性，今天被视为不可或缺的制度安排，在一百年前可能根本不存在。保留地指的是一种边界：效率论证必须在这里止步。这条边界不是空间上的线，而是一种制度行为者在面对具体决定时，主动做出的“这件事不打折”的宣告。这宣告之所以有效，不是因为它有强制力的后盾，而是因为它被足够多的人、在足够长的时间里、在足够关键的节点上，当作一件不需要论证其“成本-收益”的事来遵守。

这个宣告的发生本身就是一个值得追问的历史过程。在法律史上，“裁决者必须亲自听取被告人的最后陈述”这一要求，并不是天然就存在于所有法律传统中的。它在英国普通法中的确立，经历了漫长的演变。在中世纪的纠问式诉讼中，被告人更多是程序的客体，而非有独立陈述权的主体。直到十七、十八世纪，随着对抗制诉讼的兴起和辩护律师的广泛介入，被告人作为程序主体的地位才逐渐被制度构建起来。“最后陈述权”被确立为不可简化的仪式，不是因为某一天有人从道德哲学中推导出了这项权利，而是因为在长期的制度博弈中——法官与律师之间、王室法院与议会之间、职业法律阶层与王权之间——这项权利被一步一步地锻造为“刑事审判的构成性要素”。没有它，审判就不再是审判，即使法条上的一切都还合规。它的确立是一场发生，而不是一个发现。认识到这一点，有助于避免一种误会，即本文所说的“保留地”是在主张某种超历史的“原初道义空间”存在着、等着我们去回归。保留地不是天然存在的，它是被历史性地产出的。它的产出，需要制度行为者在具体的历史关口上做出拒绝效率论证的姿态。本文对保留地的建构性论述，不是要提出一份详尽的操作手册——那将违背道义无标度性的核心主张，因为它恰恰拒绝被写成可以被逐条落实的“标准

”——而是要辨识出那些在制度场域中效率论证最易入侵的接口，并论述在这些接口上应当维持的警觉。

第一个接口：当权力面对自己时。在刑事程序中，当控方提出一项证据，而辩方仅提出“该证据可能系非法取得”的初步线索时，控方就必须承担证明该证据合法取得的责任。在一些法域，这种证明需要达到“排除合理怀疑”的标准，并且需要经过专门的、独立的庭前审查程序。这套制度安排的实质，是国家在追诉一个人时，必须承受主动审查自己追诉行为正当性的义务。这个义务不设“在多数案件中审查没发现问题所以下次可以少审查一些”的优化空间。审查没发现问题，不是“浪费了一次审查”，而是“完成了一次审查”。这两句话的区别，就是道义无尺度带与风险管理框架的区别。在风险管理框架里，多次审查没发现问题会被解读为“审查频率可以降低”，因为风险被经验地证明是低的。在道义无尺度带里，审查没发现问题被理解为“这一次审查的完成，就是这一次的正当性得到兑现”——它不累积为未来的经验数据，因为每一次审查都是对“这一个被追诉者”的独立负责，不可由之前一百次“没发现问题”来替这一次省略。

这个不可累积性，就是无尺度性在这一接口上的制度表现。对它的侵蚀，通常不采取“从此不审查了”这种公然的方式，而采取把审查的触发门槛悄悄推高的方式。立法者不会说“非法证据排除规则废止”，他会说“辩方提出非法证据排除动议时，需要提供‘具体、明确’的线索，而不能仅凭猜测”。这个要求在法理上完全合理：不能允许辩方用毫无根据的“我猜证据是非法的”来拖延诉讼。但在实际运行中，“具体、明确”这四个字给了法官一个巨大的裁量空间，而这个裁量空间在绩效考核和办案期限的压力下，会系统性地朝着提高门槛的方向倾斜。一个需要赶在月底前结案的法官，面对辩方提出的“泛泛的”非法取证主张，驳回动议的成本远低于启动一场耗费时日的证据审查。每一次这样的微观决定，都在加固一个制度惯性：启动自我审查的门槛，在实际操作中比法条文本所显示的要高得多。当这个惯性充分形成之后，不需要任何人宣布“自我审查的义务被削弱”，审查在事实上的萎缩已经完成了。

第二个接口：当裁决者面对被裁决者时。一个法官用两个小时听取一个被告人关于其早年家庭创伤的陈述，这段陈述可能与法律上的定罪量刑要件没有直接因果关联，法官听完之后可能也没有改变判决的刑期。这两个小时在效率账本上是净亏损。然而，这两个小时在两个不同层面上划出了一条界线。在制度层面，它为“被裁决者仍然是人、不是案件编号”这个宣告提供了一个活着的载体——不是写在法条里，而是在那间法庭上，在那两个小时里，被实际地兑现了一次。在裁决者心智层面，它让裁决者在做出剥夺性决定之前，被迫在另一个人的全部叙述面前多停留两个小时。这两个小时里他也许没有形成任何新的法律判断，但他对这个人的感知发生了某种法律分析无法捕捉却真实存在的改变。这种感知改变，是判决从“严格但不公正”走向“严格且公正”的不可或缺的一步。这一步无法被管理，因为它不在绩效指标的射程之内，但你可以在制度上要求法官必须坐在那里听完。这就是道义止步之地的第二种形态：不是程序安排不能改革，而是论证负担的分配不能颠倒——不

是要求审慎证明自己“有用”，而是要求效率证明自己“确定不会在砍掉这一环节时切断裁决者与被裁决者之间那道仅存的、不可还原的感知通道”。

第三个接口：当判断面对规则时。任何成熟的法律体系都包含严格规则和一般条款两类规范。严格规则的优势在于清晰可预测，一般条款的意义在于为那些溢出类型、带着独特性的个案留下被认真对待的空间。司法改革中，“统一裁判尺度”被频繁提及，其理想是让相似案件得到相似判决，这本身是法治的核心诉求。但“统一尺度”在制度实践中有两种极为不同的实现方式。一种是通过深厚的法教义学共同体建设——法官因为接受相同的训练、浸淫在相同的解释传统中而自然地趋向类似判断。另一种是通过不断压缩自由裁量空间——细化量刑指南、统一裁判指引、将一切判断锚定在被上级法院可量化审查的“要件”上——把法官变成统一标准的执行器。后一种方式，构成了对第三种道义无标度带的侵蚀。它的逻辑是把法官的判断本身当作一种需要被管理的风险源——不同的法官会判出不同的结果，这是“不确定性”，而不确定性是必须被消除的管理缺陷。每一次对这个逻辑的落实，都显得极其合理，但累积的终点是：法官从一个拥有实践智慧的人，变成了一个标准化决策终端。而算法的判决，即使完全准确，也无法完成只有人类裁决者才能完成的那件事：对被裁决者说——“我听见你了，我犹豫过，但我最终没有被你说服。这个判决是我作为一个人做出的，我承担它的重量。”

这三种保留地——权力对自己保持怀疑、裁决者认真倾听被裁决者、判断不被算法完全替代——不是在提出三个可分别实现、分别考核的制度目标。它们根本不在绩效考核的射程之内。它们的存在方式，不是作为法律系统的“组成部分”，而是作为法律之为法律的那道门槛。这道门槛不需要被写成法条——你无法用法条写出“此处道义无标度”，就像你无法用食谱写出“此处应怀着对食材的敬意”。但你可以制度设计时，在这些接口上反复做一件事：不要在每一条效率改革的尾部，都要求那个被质疑的环节去证明自己“为什么不该被优化”。把论证负担反过来——要求效率改革者去证明“我确定这一刀不会切在标尺之外的保留地上”——这个反转本身，就是在守护那道门槛。而守护，不是回归一个想象中的原初道义状态，而是在每一个具体的、历史的制度关口上，重新做出那道“这里不让效率进”的宣告。

五、结论与证伪条件

本文从法律实证主义与自然法理论共享的一个预设出发——将法律视为已完成的对象，而后争论其道德资格——转向追问法律如何逐步丧失其道德资格。通过推进一种关于制度动力学的诊断，本文在最要害的一步退回更底层的预设：道义与效率之间的制度性博弈中，道义之所以持续失利，不是因为它分量太轻，而是因为它根本不该被放上天平。一旦它被当作“一种虽看不见但也应重视的回报”进入度量场，它就注定会在每一轮具体的制度博弈中被证明“不够确实”而败下阵来。这个持续的过程，不是道义的某种“部分损耗”，而是一个质变——当“道义可以被度量”这件事本身不再被制度行为者视为问题时，法律的道义对话能力就已经死了。

基于此，本文提出了“道义无标度性”这一概念，用以指称法律系统中某些根本性正当性要求所具有的一种性质：它们拒绝被纳入任何成本-收益核算，而正是这种拒绝构成了它们的道义内核。本文通过中国认罪认罚从宽制度中速裁程序的具体制度节点，追踪了这种无标度性如何在真实的制度环境中被系统性地解除——不是通过邪恶的法条，而是通过无数个看似合理的、以“合理配置资源”为名的微观制度决定。本文进而通过与价值理性理论、权利王牌论、不可通约性概念、卢曼系统论法学和老子退化链的对话，论证了这一概念不是已有理论的应用，而是切出了一个既有理论因其发现学惯性而未能抵达的辨别面：它指涉的不是行动者主观持有的价值信念，而是制度必须维持的一道拒绝度量的边界；它的失守，不是道义被效率击败，而是道义被移出了它唯一能存活的那个世界。

本文不声称发现了某种超历史的“道义本质”。道义无标度带的具体形态——它在哪个制度节点上被确立、以何种程序安排为载体——在不同法律传统和不同历史时期会有显著差异。本文论证的不是这些具体形态的不可改变，而是那道使它们得以可能的前提：制度必须为道义保留一个“此处不衡量效率”的空间，而不论这个空间在特定历史条件下由哪些具体安排来承载。守住这个前提，不等于守住任何一套固定的程序规则——程序可以改革、制度可以更新——但改革不能以取消“不衡量”这个姿态本身为代价。

本文的论证是脆弱且可被证伪的，因此必须在此给出明确的证伪条件。本文所立的核心判断——道义无标度性构成法律与暴力之间那条最终的界线——在以下条件下将被严重削弱或推翻。如果未来有精细的跨国司法比较研究证明，法律体系中那些被本文视为承载道义无标度性的制度安排——充分时间保障的庭前律师会见、辩方仅需提出初步线索即可触发的高成本证据审查程序、裁决者必须履行的不可简化的陈述听取义务——在被系统性地简化之后，公众对该法律体系的规范性信任和自愿服从意愿并未出现同期显著下降，且这一发现在不同法域、不同文化背景中具有高度可重复性，那么本文论证的经验基础就面临根本性的挑战。这种情形的出现将表明，本文所定义的“道义无标度性的违反”并非是公众感知法律正当性的决定性因素——决定性的因素可能在别处：整体政治制度的正当性水平、经济绩效带来的弥散性信任、或是其他本文未识别出的替代性道义承载机制。

在获得这样的反面证据之前，本文坚持它从论证中获得的核心判断。在当代司法改革的公共话语中，“提高效率”“消除拖延”“科学管理”已经日益成为无需论证的前提。本文的论证不分担这个前提，也不直接反对它，而是试图在它畅通无阻的推进路线上，设置一道它无法用自己熟悉的语言来回答的追问：你确定这一刀，没有切在标尺之外的东西上吗？

这道追问没有终极答案，但它有一个可以被实际操作的检验方式。在任何一项被提出的效率改革面前，不要求反对者去证明被砍掉的环节“有用”。反过来：让改革者证明——不是用“在大多数案件中这个环节没有发现问题”这种已经在逻辑上预设了道义可以被打折的说法，而是用比这更严格的、承认道义无标度性的标准——他想要优化的那个对象，是法律与暴力之间那条已经最细的线吗？如果他无法在给定这个标准之后仍然自信地回答“不是”，那么这场改革试图优化的，就不只是一

个程序环节。它试图优化掉的，是法律还能称其为法律的最后一层保障。

注释

[1] 本文在此处及后文中所称的“诊断”，是对近二十年来司法管理研究中一种主流观点的概括性指认。该观点强调司法效率指标的可量化性与程序审慎收益的不可见性之间的不对称，是本文论证的重要对话起点之一，但并不特指某一位学者的体系。本文对这一诊断的充分承认同时伴随根本性的不满，详见后文。

[2] “程序性奢侈”是本文对一类强调程序冗余之不可化约价值的理论立场的概括性命名，用以指代上述诊断中那些将程序审慎视为“奢侈”而非“构成性要求”的论述。本文与此种立场共享某些关切，但在最底层的预设上分道扬镳，详见本节。

[3] 本文与韦伯的对话主要立足于其《经济与社会》中所确立的这对理想类型。韦伯的理论焦点在于行动者主观意义层面；本文所提出的“道义无标度性”则试图在这一经典区分之上切出一个制度性维度——即制度维持拒绝度量之“姿态”的边界，它独立于行动者主观持有何种价值理性。

[4] 德沃金在《认真对待权利》中将个人权利喻为“王牌”，其理论重在论证权利在政治道德论辩中的优先地位。本文并不否认权利优先，但指出“王牌”隐喻仍然预设了同一天平上的优先考量，而“道义无标度性”所要拒绝的恰是权利被放上天平进行权衡这件事本身。故两者的分歧不在于“权利有多重”，而在于“权衡”这一姿态是否应被允许进入某些领域。

[5] 本文对卢曼的对话主要针对其《社会中的法律》一书中提出的“法律是操作上封闭、认知上开放的自创生系统”这一核心命题。卢曼的理论旨在抵御政治、经济、道德等系统对法律的过度干预，却未充分应对法律系统内部主动引入效率管理技术而导致的“自我殖民”现象。本文的诊断正切在此一直区。

[6] 材料说明：本节所使用的制度描述、典型场景及时间数字（如“不足三十分钟”），系综合公开的司法改革文件、既有研究中的观察以及实务讨论中的常见情形而提炼的例示性材料。为配合论证需要，已进行典型化、场景简化与匿名化处理，旨在展示一种制度趋势的发生学逻辑，而非对特定法院、特定时段精确统计数据的呈现。读者不宜将其视作经过同行评议的经验证据，也不应将其直接等同于某一具体司法案件的客观记录。

[7] 本文借取老子《道德经》“失道而后德……”的退化链意象作为宏观参照框架，旨在描述规范化秩序依序丧失内在自发性的现象学过程。这一借用并不表示本文全盘接受道家哲学的形而上学预设，亦非旨在对老子文本进行传统的经学阐释，而是将其作为理解法律道义根基塌缩的一种思想坐标。

参考文献

Berlin, Isaiah. 1969. Four Essays on Liberty.

Dworkin, Ronald. 1977. Taking Rights Seriously.

Luhmann, Niklas. 1993. Das Recht der Gesellschaft.

Raz, Joseph. 1986. The Morality of Freedom.

Weber, Max. 1922. Wirtschaft und Gesellschaft.

老子. 《道德经》.