

从要件识别到变量评估： 刑事司法如何破解“定性难题”

桑本谦*

目次

一、问题的界定

二、变量评估与定罪量刑

(一) 描述

(二) 解释

三、演化分析

(一) 法律与道德直觉

(二) 法律与法学知识

四、结语

摘要 “法律先于法学”的事实意味着负责罪责评价的人类道德直觉是刑法最初的立法依据，而刑法理论回避解码道德直觉，选择了“填充犯罪分类表”的捷径，虽有刑法教义作补充，但终究没能揭示刑法的底层逻辑。罪责评价的绝对尺度是犯罪的预期损失，因此定罪量刑需要考虑破案率、作案成功率等决定刑罚威慑水平和犯罪预期损失的各种变量。以变量评估取代要件识别，可以有效应对刑事司法实践中的定性难题。

关键词 犯罪构成理论 犯罪构成要件 抢劫 诈骗 边际威慑

一、问题的界定

如果某个违法行为既像此罪又像彼罪，则即使案件事实清楚、证据确凿，司法机关也可能拿不准该以何种罪名立案、起诉或定罪，此时司法实践就会遭遇“定性难题”。^{〔1〕} 除了在此罪和彼罪之间模棱两可，“定性难题”有时还会表现为在出罪和入罪之间左右为难。譬如侵权和犯罪之间就不是泾渭分明的，虽然广义上的侵权包括所有的犯罪，但具体到某个侵权行为是否构成犯罪，司法机关内部也屡有分歧。以下几个真实案例可以让我们对“定性难题”有个直观的了解。^{〔2〕}

* 中国海洋大学法学院教授、法学博士。

〔1〕 这里所说的定性难题不包括“法条竞合”的情形，同样是一个违法行为符合两种犯罪的构成要件，但后者不存在定性难题。

〔2〕 这几个案例都是检察机关公诉部门提供的，情节略有删减。

案例1: 盗窃还是侵占?

犯罪嫌疑人洪某为某超市雇员,负责卖货收银,该超市在工商局登记为个体工商户。在长达3个半月的时间里,洪某利用职务之便,采取虚增电子支付收入(微信、支付宝、银联)套取现金的方式非法获利8万余元。该案破获后,检方拟以盗窃罪起诉,但有异议认为该案构成侵占罪。

案例2: 诈骗还是无权处分?

犯罪嫌疑人黄某谎称其受某市烟草局委托管理一片海水养殖区域,受害人信以为真,遂同意与黄某签订租赁合同,黄某因此非法获利3万元。检方拟以诈骗罪起诉,但有异议认为主张该行为不构成犯罪,属于民法上的无权处分。

案例3: 抢劫还是寻衅滋事?

某日凌晨1点左右,犯罪嫌疑人蓝某在某餐馆吃饭后拒付餐费,受害人(餐馆老板)发现后便尾随其后。蓝某打算逃跑又恐受害人穷追不舍,遂用酒瓶猛击其头部后逃出餐馆。受害人被击倒后立即起身紧追,跑出约800米后将犯罪嫌疑人追获并制服,还夺下了其手中一把水果刀。随后报警,警方赶赴现场后将犯罪嫌疑人抓获。经法医鉴定,受害人头部构成轻微伤。检方拟以抢劫罪起诉,但有异议主张构成寻衅滋事罪,且情节显著轻微,不构成犯罪,可移交公安机关以治安管理处罚。

这几起案件都不复杂,司法机关的纠结关键在于“定性难题”。案例1中的犯罪行为既像盗窃又像侵占,犯罪嫌疑人作为超市雇员负责卖货收银,此工作性质算不算为受害人“代管财物”就成了争点,这是区分盗窃和侵占的法定要件。案例2中的违法行为当然属于无权处分,但同时还完全满足构成诈骗罪的所有法定要件,司法机关如何定性事关行为人人罪出罪。案例3中的违法行为确实具备抢劫罪的外观,以暴力取财和以暴力逃债理论上没什么不同,但该行为也同时满足构成寻衅滋事罪的法定要件,该案的定性同样事关入罪和出罪,只要被定性为寻衅滋事,以“情节显著轻微”论处就不算过分。

为何司法实践屡屡遭遇“定性难题”?说来话长,但若追根溯源,与其说问题出在实践环节,倒不如说指导实践的理论本身就先天不足。通过对犯罪行为分门别类,进而归纳总结不同犯罪的具体特征,传统刑法学创造出的一套以构成要件为核心的理论体系。在这个理论体系中,每一种犯罪各自对应着一组构成要件,因而刑事司法实践的核心技术就是“要件识别”——面对一种具体的违法行为,司法机关先根据法定要件去判断这种行为是违约、侵权、行政违法还是刑事犯罪;构成犯罪的,则同样需要根据法定要件去判断该行为构成哪种犯罪。尽管这种司法模式可以顺利解决大部分案件,但若一个违法行为同时符合或贴近两种犯罪的构成要件,就难免出现“要件失灵”,定性难题正是要件失灵的结果。

要件失灵了,不是还有教义吗?话虽这么说,传统刑法学也确实含蓄地承诺,刑法教义足以应对要件失灵;甚至为了凸显教义的功能和地位,还干脆用“教义”冠名了理论——不过老实说,“刑法要件学”要比“刑法教义学”更加名副其实。且不说教义本身的实战能力尚待严格检验,即使教义确实有用,在其供给和需求之间也存在大量缺口。就处理前文所列的那几个案子而言,刑法学教义又在哪里呢?总不至于把“刑法公正”“校正正义”或“罪刑相适应”之类的“终极杀器”祭出来吧?祭出来又有何用?

缺乏坚实的底层逻辑,任何刑法学理论在应对争议案件时都难免尴尬处境。必须承认,没有哪种法学理论可以消除司法实践中所有争议,毕竟法律问题涉及太多的主观评估,不可能有足够的据来矫正并统一人们的认识,但是,争议的底线和范围仍可作为评价理论优劣的重

要指标。^{〔3〕}如果争议的范围太广，且底线太低，那么理论指引司法实践的能力就难免遭到质疑。坦率地说，笔者之所以怀疑传统刑法学理论的功能和质量，不是因为前文所列的案例引发了争议，而是因为这些争议实在是档次不够。

不仅争议的级别太低，争议的质量也不高。争议双方使用同一套理论支配的技术流程，却并未形成真正的交锋。正反双方都用要件识别的套路来论证自己的观点，但在相互反驳时却都说不出个所以然。有争议很正常，真理越辩越明，但若每个人都觉得真理站在自己一方，却不能解释别人为什么错了以及错在哪里，辩论就形成不了共识，争议的解决最终还要依靠某种权力因素。

这是否意味着传统刑法学理论本身缺乏论证的力量，无力把争议双方收敛或聚合到同一个轨道？无论确认还是排除这种可能性，都需要一个理论比较的视野。对于刑事司法如何破解定性难题，本文主张从“要件识别”转向“变量评估”。一旦要件失灵，就无须在要件含义或行为性质上继续纠结，而应该探寻隐藏于要件背后的底层刑法逻辑，并在此基础上，评估决定犯罪预期损失和刑罚威慑水平的各种变量，尤其是破案率和作案成功率，进而根据评估结果去定罪量刑。从要件识别到变量评估，当然会牵涉理论视野的切换，法律经济学是变量评估的理论支撑。

但这个主张肯定会招致各种质疑。法律人会说，破案率和作案成功率从来不是刑法上任何犯罪的构成要件，它们至多是量刑的酌定情节，无论如何也不可能影响定罪。更何况，没有任何证据显示立法者在制定刑法时，慎重考虑了破案率和作案成功率的因素，倘将其作为定罪量刑的依据，就难免有主观臆断之嫌。

为了澄清上述质疑，本文将在第二部分论证：破案率和作案成功率既可以是量刑的基础，又可以是定罪的依据。作为刑事法律制度的两个内生变量，破案率和作案成功率在刑法制定之初就已渗透进犯罪构成和刑法教义之中。第三部分致力于从历史维度揭示破案率和作案成功率进入犯罪构成和刑法教义的过程，涉及刑事法律制度的起源和演变，且在解码道德直觉之余，还会对传统刑法学理论做一番知识考古。

二、变量评估与定罪量刑

发生于2006年的许霆案曾轰动一时，该案一审以盗窃金融机构罪判处许霆无期徒刑，不仅引爆了舆论，而且引发了大量学术讨论。但对于一审判决量刑畸重，公众、媒体以及法律人等各个圈层都不缺少共识，难题在于如何说理。为了给许霆开脱罪责，法律人提供的理由五花八门，却无一条经得起推敲。^{〔4〕}该案二审最终适用《刑法》第63条“特殊减轻”改判许霆五年有期徒刑，其依据却仅仅是权威学者提出的许霆“主观恶性不大”“违法程度较轻”“责任程度较轻”以及“期待可能性

〔3〕 科学是规范认知的一个制度性体系，其特点是具有征服反对性见解的力量。总体来说，越是接近于科学的领域，其争论就会基于越多的共识。所以波斯纳认为，科学趋向于合流，而道德哲学趋向于分流。参见[美]波斯纳：《道德与法律理论的疑问》，苏力译，中国政法大学出版社2001年版，第60页。

〔4〕 为了给许霆开脱罪责，法律专家提供了各种理由。包括：（1）银行有过错，甚至认为银行有引诱犯罪之嫌，更甚者认为银行应该向许霆道歉；（2）金融机构不应受到特殊保护，因为“法律面前物物平等”；（3）自动取款机未必属于金融机构；（4）司法应该告别“算数司法”和“机械司法”等等。关于这些理由的分析和批评，参见桑本谦：《传统刑法学理论的尴尬——面对许霆案》，载《广东商学院学报》2009年第5期，第67—74页；桑本谦：《理论法学的迷雾》，法律出版社2015年版，第109—119页。

有所降低”等表面化的理由。^{〔5〕} 在传统刑法学的理论框架中,说理和论证也只好到此为止。这起被称为“教科书式”的轰动案例,其实并未真正发挥教科书的作用。

套路化的措辞抓不住要点(这些措辞差不多可以适用于所有从轻或减轻处罚的场合),要为许霆开脱罪责讲出过硬的理由,就必须分析许霆的作案方式。首先,许霆作案成功纯属巧合,利用自动取款机在升级过程中发生故障而恶意支取现金,虽被定性为盗窃金融机构,但事先看来,这种作案方式的成功率几乎为零。其次,许霆使用自己的银行卡和自己设定的密码并在监控录像下完成作案,这种作案方式的破案率接近百分之百。实际上,在警方介入之前,商业银行通过查询账户就已经破获了案件,精准锁定了犯罪嫌疑人。而如果许霆按银行通知的要求将现金悉数返还,银行就不会报案,这起轰动案件也就根本不会发生。

那么,“作案成功率几乎为零”且“破案率接近百分之百”意味着什么? 首先,犯罪的预期损失相当于犯罪的实际损失和作案成功率的乘积,所以,如果一种犯罪的作案成功率极低,就意味着这种犯罪的预期损失很小,即使减轻甚或免于处罚,也不至于导致犯罪失控。^{〔6〕} 其次,刑罚的威慑效果相当于惩罚严厉程度和破案率的乘积,所以如果一种犯罪的破案率很高,则无需严厉的惩罚就足以阻止犯罪。^{〔7〕} 综合以上两个理由,就足以解释为什么许霆案可以在法定刑以下减轻处罚。这起案件非同寻常,如果盗窃金融机构的罪犯都采用许霆的作案方式,则不仅司法机关额手称庆,金融机构也可以高枕无忧了。刑法专家之所以认定许霆“主观恶性不大”“违法程度较轻”以及“责任程度较轻”,其深层依据就是许霆的作案方式显著区别于同类犯罪。在做出如上判断之前,刑法专家的道德直觉已经下意识地评估了这种特殊作案方式的破案率和作案成功率,但在“知其然而不知其所以然”的情况下,刑法专家没办法把其中的道理表达出来。

(一) 描述

变量评估在许霆案中的应用价值如何? 貌似可以影响量刑,但还不至于影响定罪,毕竟这两个变量从来不是法定的构成要件,甚至不是法定的量刑情节。但其实不然,定罪和量刑之间并没有严格的界限。刑法之所以需要把犯罪分门别类,就是为了设计出有针对性的量刑方案,如果所有犯罪的量刑方案完全一致,那么理论上就没有区分犯罪的必要。试想,倘若在修订刑法时,立法者把许霆的作案方式——利用自动取款机的故障恶意支取现金——从盗窃金融机构中单列出来,另立新罪,那么破案率和作案成功率就从仅仅影响量刑变成了可以直接影响定罪。刑法之所以没有为此另起炉灶,只是因为这种作案方式太罕见了,千年等一回,与其另立新罪,还不如经最高法院核准更能节省刑事司法的制度成本。^{〔8〕}

但把侵占从盗窃中分离出来就意义重大了。盗窃显然是比侵占更古老的犯罪,起初侵占是被

〔5〕 参见张明楷:《许霆案的定罪与量刑》,以及陈兴良:《许霆案的法理分析》,两文均载《人民法院报》2008年4月1日,第005版。

〔6〕 关于作案成功率和犯罪实际损失如何影响犯罪预期损失的函数关系, see Steven Shavell, *Deterrence and the Punishment of Attempts*, 19 *The Journal of Legal Studies* 435-466 (1990).

〔7〕 关于惩罚概率和惩罚严厉程度如何影响刑罚威慑效果的函数关系, see Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 *Journal of Political Economy* 169-217 (1968). 本文所说的“破案率”只是惩罚概率的近似值。

〔8〕 这里涉及两种成本之间的权衡,另立新罪可以提高法律的精确性,但会产生修改法律和实施法律的制度成本;但不立新罪须经最高法院核准,这也会产生额外的司法成本。关于法律精确性和实施成本之间利弊权衡的经济学逻辑, see Louis Kaplow, *Economic Analysis of Civil Procedure*, 23 *The Journal of Legal Studies* 307-401 (1994).

包括在盗窃之中的，^{〔9〕}直到法律决策者发现，作为盗窃罪的一个子集，侵占是非常特殊的，无论其犯罪嫌疑人为受害人代管财物，还是持有受害人的遗忘物或埋藏物，都很容易被锁定，破案率接近百分之百，通常受害人依靠自己的力量就能破获案件，所以决定将这类特殊的盗窃单列出来，另立侵占罪名。难怪在汉语中“特殊”和“另类”是同义词。但是请注意，将侵占从盗窃中分离出来的原因，不是犯罪嫌疑人为受害人代管财物或持有受害人的遗忘物或埋藏物，而是因为这类犯罪的破案率接近百分之百，受害人无须警方的介入就可自破案件，因此可以适用自诉程序。

既然如此，刑法为何不把“破案率极高”直接规定为构成侵占罪的法定要件？答案当然是操作性障碍。法官自由裁量权若被过度扩张，各种风险都很大——司法的难度和成本上升，法律的激励信号模糊不清，枉法裁判和徇私舞弊也会乘虚而入。^{〔10〕}为了解决操作性障碍，立法者只能将实质性要件形式化处理，“破案率极高”因此被具体化为几种特定情形——代管受害人财物、持有受害人的遗忘物或埋藏物等等。但这只是侵占最常见的表现，列举不能穷尽，侵占罪的外延一直是开放的。不仅如此，形式化立法的先天不足使法律因此趋向于破碎化，被要件割裂之后的法律难免产生缝隙和重叠，只要某个违法行为恰好发生于法律的缝隙或重叠地带，要件失灵就会引出定性难题。一旦要件失灵，继续纠结要件就没什么意义了。司法机关应该回到原始的实质性思路，需要分析作案方式并预测犯罪后果，而评估破案率和作案成功率则是其中的关键。

在前文列举的案例1中，司法机关的内部争议集中于，犯罪嫌疑人作为超市收银员，其工作性质算不算为“受害人代管财物”，但在要件失灵的条件下拿法律概念来做文章是注定无果的，回到原始的实质性思路需要评估破案率。至于破案率高到何种程度才构成侵占？这个问题不难回答。司法机关只要懂得——正是由于侵占罪的破案率远高于盗窃罪，刑法才对前者规定了很轻的处罚；并且正是由于受害人无须借助警方的力量而自破案件，刑法才将侵占列入告诉才处理的自诉案件——答案就显而易见了：当破案率高到受害人只需仅靠自己的力量就能大致准确地锁定犯罪嫌疑人的地步，就可以认定为侵占。根据这个逻辑，案例1的关键，就不再是犯罪嫌疑人作为超市收银员的工作性质，而是受害人雇佣的收银员数量。如果受害人只雇佣了一名收银员，那么当雇主发现超市现金账目出现异常时，就可以基本确定犯罪嫌疑人非此人莫属。

刑法上的自诉案件兼具侵权和犯罪的特征，实际上处于两者之间的过渡地带。当司法机关难以区分一种违法行为是犯罪还是侵权的时候，同样可以通过评估违法行为的破案率来解决定性难题。如果破案率逼近甚或达到百分之百，通常可以认定为侵权而非犯罪，因为破案率是区分侵权和犯罪的一个重要变量。如果一种违法行为的破案率逼近或达到百分之百，无须警方介入受害人就能自破案件，法律就可以将起诉和举证的任务全部交给受害人承担，国家的职责只是提供一个裁断是非的法院，这种处理案件的方式就是民事司法。相反，如果破案率很低，非警方介入不可，则立案、侦查、起诉、举证、审判以及强制执行等各个环节的责任就要全部由国家承担，这种处理案件的方式就是刑事司法。从法律经济学的角度，法律之所以需要区分犯罪和侵权，并设置两套不同的司法程序，就是为了发挥私人和国家各自在社会控制方面的比较优势，从而将社会控制总成

〔9〕 虽然侵占罪萌芽于罗马法，但直到19世纪才与盗窃罪分离。在中国古代，侵占一直被包含在盗窃罪之中，直到1911年《大清新刑律》才正式规定了侵占的罪名。

〔10〕 关于法律的形式合理性的经典论述，参见[美]富勒：《法律的道德性》，郑戈译，商务印书馆2007年版，第40—111页。其经济学逻辑，see Louis Kaplow, *supra* note [8], at 307-401；桑本谦：《法理学主题的经济重述》，载《法商研究》2011年第2期，第27—35页。

本最小化。^{〔11〕} 这是民刑分界的主要原因,并且这种解说远比传统法学上的“民刑分界论”更有说服力。^{〔12〕}

案例2中的违法行为当然属于无权处分,至于是否构成诈骗,则需另当别论。虽然刑法规定的诈骗罪要件包括“虚构事实或隐瞒真相”,但却不是所有虚构事实和隐瞒真相的行为都构成诈骗。生活中的绝大多数谎言都是合法的,交易双方在讨价还价的过程中也经常信口开河,非但不会构成犯罪,有时甚至连违约或侵权都算不上。^{〔13〕} 虽说诈骗罪要件中还有“以非法占有为目的”,但虚构事实和隐瞒真相通常都有“占有”的目的,至于占有是否“非法”,这本身就涉及定性难题,且包含循环论证。当要件失灵时,评估破案率可以帮助我们区分欺诈和诈骗。如果破案率接近百分之百,受害人可以自破案件,就通常可以定性为民事欺诈;而如果破案率很低,非警方介入不可,就只能定性为诈骗罪。

当然,通过评估破案率来区分欺诈和诈骗并不总是灵验。明知自己无力偿还却通过虚构事实或隐瞒真相的手段骗取他人借款的行为,即使破案率高达百分之百,也仍可以被定性为刑事诈骗,而不只是民事欺诈。这里的关键在于,虽然受害人有能力自破案件,但却不能通过民事赔偿追回全部损失。刑事制裁是阻止违法行为的最后法律手段,如果侵权法能够内化所有的违法损失,理论上刑法就没有存在的必要。^{〔14〕} 但由于种种障碍,侵权法无力内化所有的违法损失,行为人缺乏偿还能力是破案率太低之外的另一个障碍,对于明知自己缺乏偿还能力而实施的欺诈,只有刑事制裁才能产生足够的威慑。但案例2中的违法行为则完全不属于上述情形。只要烟草局发现其海滩用地被他人处分,就可以自破案件而无须警方介入,破案率逼近百分之百,且在正常情况下,两个受害人(承租人未必是)都可以通过民事诉讼追回全部损失。因而无论在何种意义上,该案都无须以诈骗罪论处。

除破案率之外,刑事司法实践还应关注作案成功率。由于犯罪的预期损失相当于犯罪一旦成功造成的实际损失与作案成功率的乘积,所以作案成功率影响罪行轻重——作案成功率越高,罪行就越重,惩罚也越重,反之亦然。企图用巫术杀人之所以不构成任何犯罪,就是因为这种作案方式没有成功的可能。^{〔15〕} 此外,很多犯罪都以“国家工作人员”的特定身份为要件,是因为如果没有国家工作人员的身份,其作案成功率就几乎为零。

法律经济学视野中的犯罪类似于一种职业选择——一个人之所以选择犯罪,是因为在他看

〔11〕 以下文献为区分私法和公法提供了基本的经济学框架,Gray Becker & George J. Stigler, *Law Enforcement, Malfeasance, and Compensation of Enforcers*, 3 *The Journal of Legal Studies* 1 (1974); Donald C. Keenan & Paul H. Rubin, *Criminal Violations and Civil Violations*, 11 *The Journal of Legal Studies* 365 - 377 (1974); William M. Landes & Richard A. Posner, *The Private Enforcement of Law*, 4 *The Journal of Legal Studies* 1 - 46 (1975); A. Mitchell Polinsky, *Private versus Public Enforcement of Fines*, 9 *The Journal of Legal Studies* 105 - 127 (1980).

〔12〕 传统法学上民刑分界有很多论点,一个比较全面的介绍,参见于改之:《刑事犯罪与民事不法的分界:以美国法处理藐视法庭行为为例的分析》,载《中外法学》2007年第5期,第85—97页。

〔13〕 广告法会区分虚假广告和夸张描述(例如欧莱雅的广告语:“你值得拥有!”),尽管严格意义上两者都是谎言。合同法上关于赠与承诺撤销权的规定,也同样意味着撒谎或信口开河并不承担法律责任。

〔14〕 参见[美]大卫·弗里德曼:《经济学语境下的法律规则》,杨欣欣译,法律出版社2004年版,第341—361页;William M. Landes & Richard A. Posner, *supra* note〔11〕, at 1 - 46.

〔15〕 这种情形在传统刑法学理论上被称为“行为不能犯”,其经济学解释,see Steven Shavell, *supra* note〔6〕, at 435 - 466.

来,犯罪比从事其他合法职业能带来更高的收入。^[16] 作案成功率意味着犯罪如同其他职业一样都是有门槛的,如果某种身份是越过犯罪门槛的必要条件,那么把这种特定身份规定为犯罪的构成要件,司法机关就可以轻而易举地把绝大多数人排除在犯罪嫌疑人的范围之外。有时年龄就是一道犯罪门槛,有很多犯罪是未成年人望尘莫及的,法律之所以对16岁以下的未成年人豁免大部分刑事责任,除了那些众所周知的原因,还因为绝大多数犯罪门槛是未成年人无法跨越的。但暴力犯罪的门槛较低,所以当14~16岁的未成年人从事暴力犯罪的时候,刑罚就必须及时跟进。

很多刑法条文、教义和构成要件隐含了对破案率和作案成功率的双重评估。刑法对蓄谋犯罪的处罚之所以比对冲动犯罪更加严厉,主要是因为前者的作案成功率更高且破案率更低;^[17] 刑法对又聋又哑的人犯罪之所以比对普通人犯罪的处罚更轻,主要是因为前者犯罪的作案成功率更低而破案率更高。类似的情形是,相对于普通罪犯,累犯和惯犯因为犯罪经验丰富、作案技术高超所以其作案成功率较高而破案率较低,这是对累犯从重处罚的许多原因之一。^[18] “人身危险性”和“主观恶性”的概念也隐含了对破案率和作案成功率的评估。人身危险性更大的罪犯,通常是作案成功率更高、破案率更低且一旦作案成功就会造成更严重后果的罪犯;“主观恶性”的概念通常被用于衡量罪犯的犯罪意图,那些犯罪意图更强烈的罪犯,愿意为犯罪支付更高的成本,因此其作案成功率更高且破案率更低。受害人过错会提高罪犯的作案成功率或者降低其破案率,因而只要受害人能够避免过错,作案成功率就会降低且破案率会提高,这是受害人过错之所以能成为对罪犯从轻或减轻处罚的理由之一。

(二) 解释

前文的分析表明:早在刑法诞生之时,破案率和作案成功率就已渗透进犯罪构成和刑法教义之中,成为刑法的两个内生变量。但要进一步论证这个观点,就涉及刑法的整体架构——预设目标和技术路线。为了实现预设的目标,制度设计要努力探索出最优的技术路线。

刑法的主要功能是威慑犯罪(deterrence),即所谓“惩一奸之罪,而止境内之邪”,^[19] 通过少量惩罚释放威慑信号,就能打消潜在罪犯的犯罪动机;^[20] 刑法的次要功能是剥夺犯罪能力(incapacitation),被处以监禁或死刑的罪犯,即使还想继续犯罪,也是有心无力了。^[21] 威慑的对象是不特定的潜在罪犯,剥夺犯罪能力的对象是已被处罚的特定罪犯。但无论是威慑还是剥夺犯罪能力,都是为了阻止未来的犯罪。已经发生的犯罪,其损失属于沉没成本。

尽管提高威慑水平可以减少犯罪的数量,但威慑不是无成本的,无论提高惩罚的严厉程度,还是提高破案率,都要增加威慑的社会支出。两种成本此消彼长,因而最优威慑的均衡点定位于两

[16] 贝克尔最早表达了惩罚相当于给犯罪定价的观念, see Gary S. Becker, *supra* note [7], at 176-178.

[17] Richard A. Posner, *An Economic Theory of the Criminal Law*, 85 Columbia Law Review 1193-1231 (1985).

[18] 关于累犯应被从重处罚的其他原因,参见[美]波斯纳:《法律的经济分析》,蒋兆康译,法律出版社2007年版,第325—326页。

[19] 《韩非子·六反》。

[20] 关于威慑理论的开创新文献, see Gary S. Becker, *supra* note [7]. 后续的威慑感知理论, see W. William Minor & Joseph Harry, *Deterrent and Experiential Effects in Perceptual Deterrence Research: a Replication and Extension*, 19 Journal of Research in crime and Delinquency 190-203 (1982).

[21] 关于剥夺犯罪能力的最优模型, see Steven Shavell, *A Model of Optimal Incapacitation*, 77 The American Economic Review 107-110 (1987).

种成本在边际上相等的位置。^{〔22〕}但这只是泛泛而论,要实现最优威慑的宏观目标,刑法的技术路线是对犯罪分门别类,重罪重罚,轻罪轻罚。^{〔23〕}但如何衡量罪行的轻重却需要清晰的尺度并涉及一套精致的算法。

刑法是一个面向未来的激励机制,虽然惩罚发生于犯罪之后,但阻止犯罪(无论是威慑还是剥夺犯罪能力)的力量却是作用于犯罪之前而非犯罪之后。在潜在罪犯看来,定罪量刑是个概率性事件,因而刑罚的威慑效果相当于惩罚严厉程度和惩罚概率的乘积。^{〔24〕}而在立法者看来,作案成功也是个概率性事件,因而犯罪的预期损失相当于犯罪的实际损失和作案成功率的乘积。^{〔25〕}

只要分清事先和事后,就得承认定罪量刑的基础是犯罪的预期损失而非实际损失,之所以需要考虑实际损失,只是因为后者是影响前者的一个变量。^{〔26〕}风险不见得会导致损失,但风险一定会产生预期损失。一种行为只要事先看来足够危险,即使事后没有造成任何实际损失,也可能被视为犯罪,比如未遂犯或危险犯;^{〔27〕}相反,一种行为哪怕事后看来损失惨重,只要发生概率很低,也可能不被视为犯罪或只被视为轻微的犯罪,比如过失犯。两种极端情形都可以统一在犯罪预期损失的尺度上,作案成功率从趋近于零到趋近于百分之百就界定了一个可过渡的频谱。

因而所谓“罪刑相适应”,更准确的表达应该是“刑罚威慑水平与犯罪预期损失相适应”,但当我们说“罪刑相适应”的时候,意味着将惩罚严厉程度这一变量从威慑水平中分解出来,于是作为决定威慑水平的另一个变量——破案率——就被用去评估罪行轻重了,这种变换相当于分子上一个乘数变成了分母上的一个除数,可以解释为什么越是难以侦破的案件或难以被抓获的罪犯被认为罪刑越重。^{〔28〕}

罪刑相适应原则有两个功能,一是在预算约束下最大化威慑的产出(相当于犯罪的减量),这

〔22〕 关于最优威慑的非技术的介绍,参见[美] 罗伯特·考特、托马斯·尤伦:《法和经济学》,施少华等译,上海财经大学出版社2002年版,第384—390页。

〔23〕 关于跨类别的犯罪的边际威慑,不同于Shavell的模型,see Louis L. Wilde, *Criminal Choice, Nonmonetary Sanctions, and Marginal Deterrence: A Normative Analysis*, 12 International Review of Law and Economics 333-344 (1992).

〔24〕 贝克尔的模型描述了惩罚概率、惩罚严厉程度和刑罚威慑效果之间的函数关系,see Gary S. Becker, *supra* note〔7〕, at 172-180.

〔25〕 《刑法》第61条关于量刑的一般原则规定,量刑需要综合考虑“犯罪的事实、犯罪的性质、情节和对于社会的危害程度”,尽管这个表达由于概念之间的交叉重叠而显得有些模糊,但至少明确了,量刑不能单纯考虑犯罪的实际损失。

〔26〕 计算犯罪的预期损失,还需要减去犯罪的产出,许多犯罪是有产出的,比如复仇会威慑挑衅,诈骗会减少愚蠢,如果产出足够大就会产生犯罪的阻却事由,比如紧急避险或正当防卫。此外,在贝克尔的意义上,犯罪所得也是犯罪产出的组成部分,比如盗窃所导致的社会成本主要是防控犯罪的社会成本,而犯罪所得不属于社会成本,而只是受害人的私人成本。社会成本是社会财富的真正减少,私人成本只是社会财富的转移。

〔27〕 为什么有些危险行为属于犯罪而有些危险行为只算是侵权? 一个非常重要的区分标准仍然是破案率。私藏枪支是犯罪,而私养一只老虎却是侵权,根本区别不是危险和伤害的大小,而是前者很隐蔽,而后者却是暴露的,无需警方介入,受害人(通常是邻居)通过民事诉讼就可以排除妨害。如果一种行为一旦引发事故就会损失惨重,但只要事故概率足够低,也可能不被视为犯罪,或只构成轻微犯罪,前者如建造核电站,后者如过失犯罪。关于区分危险犯和危险行为的理论分析,see Steven Shavell, *supra* note〔6〕, at 435-466.

〔28〕 设惩罚的严厉程度(实际损失)为Y,惩罚概率为P1,犯罪的实际损失为X,作案成功率为P2,则“罪行相适应”可以被描述为 $Y = f(X)$ 的函数,在一定区间内是个正比例函数。令 $YPI = XP2$,则 $Y = \frac{XP2}{PI}$,但实际上,由于死刑封顶,惩罚的严厉程度不可能与犯罪的预期损失呈线性增长,因此 $Y = f(X)$ 在平行于横轴之前更接近于一个对数函数。当然目前还没有任何一个数学函数可以精准描述罪与罚之间的正向关系。

类似于农民根据土地状况和农作物价值在不同地块上投入不同数量的化肥，以最大化农作物产出。二是使刑法保持一种边际威慑——这是一种以轻罪取代重罪的激励，可以让犯罪更加理性，防止“一不做二不休”。〔29〕

如果抢劫和盗窃都可以实现“非法占有”的犯罪目的，那么刑法对抢劫规定更重的处罚，就可以激励潜在罪犯尽可能选择盗窃而不是抢劫来实现犯罪目的。虽然抢劫给受害人造成的实际损失未必比盗窃更大，但之所以会受到更加严厉的处罚，是因为抢劫会伴随着人身伤害的危险，增加的处罚对应于人身伤害的预期损失。入户盗窃之所以比一般盗窃的处罚更重，是因为入户盗窃比一般盗窃更可能转化为抢劫，增加的处罚对应于盗窃转化为抢劫而增加的人身伤害的预期损失。入户抢劫之所以比一般抢劫的处罚更重，是因为入户抢劫的作案成功率更高，且给受害人造成更大的恐惧和更高的人身伤害风险。受害人面对抢劫只有四个选项：一是对抗，二是不作为，三是顺从，四是逃跑。只有逃跑成功，受害人才可免于伤害。而面对入户抢劫，受害人几乎完全丧失了逃跑的选项。受害人面对抢夺的状况就好得多，只要不选择对抗，就至少可以免于人身伤害。〔30〕

回到案例3，虽然行为人确实造成了人身伤害，但事先看来，这种违法行为给受害人造成人身伤害的风险却远低于一般的抢劫。行为人的目标明确，仅限于强吃“霸王餐”，受害人很容易避免人身伤害。但抢劫的受害人却要面对完全不同的风险：即使乖乖地把财产交出来，也不见得就能避免人身伤害，因为抢劫的目标是未知的，且存在信息不对称——即使受害人已经交出全部财产，抢劫犯仍可能误以为使用暴力还能获得更多的财产。虽然案例3中的违法行为完全满足抢劫罪的所有构成要件，但面对抢劫和强吃“霸王餐”，受害人控制人身伤害风险的能力截然不同。就作案目标明确而论，强吃“霸王餐”的行为更接近于抢夺，事后的暴力事件则纯属偶然。

刑法规定，犯盗窃、诈骗或抢夺的，如遇受害人反抗而使用了暴力，就会从原罪转化为抢劫，并比照抢劫定罪量刑（《刑法》第269条）。但从前文的分析可以看出，从受害人控制人身伤害风险的尺度来看，抢劫不同于转化型的抢劫，抢劫受害人的的人身伤害风险在事先看来不仅是完全未知的，而且是难以控制的；而转化型的抢劫中，受害人的的人身伤害风险至少在转化之前是可控的，或者说，受害人有能力控制原罪转化为抢劫的风险。在这个意义上，对于转化型的抢劫，与其以抢劫罪论处，不如以原罪加故意伤害罪数罪并罚。

上述分析表明，受害人控制伤害风险的能力可以影响犯罪预期损失，也可以影响作案成功率，并最终影响定罪量刑。法律对蓄谋犯罪的处罚之所以比对冲动犯罪的处罚更重，除了前文讨论的原因之外，还因为面对蓄谋犯罪，潜在受害人很难控制风险。但面对冲动犯罪就完全不同，冲动犯罪通常涉及受害人的挑衅，因此只要潜在受害人约束自己的行为，就能有效控制伤害风险。

在某种意义上，“即兴犯罪”比蓄谋犯罪和冲动犯罪更加危险。曾轰动一时的药家鑫案就属于“即兴杀人”。受害人被药家鑫撞伤之后只是记下了他的车牌号，绝不可能预料到这种正常反应居然会被杀人灭口。一个人要想活得安全，可以不惹事、不生非，小心谨慎约束自己的行为，以避免卷入某个仇杀或情杀。但若遇到药家鑫这种罪犯，这些防范措施就会全部失灵，死神从天而降，事

〔29〕 斯蒂格勒最早提出“边际威慑的概念”，see George J. Stigler, *The Optimum Enforcement of Laws*, 78 *Journal of Political Economy* 526 - 536 (1970). 此外，沙维尔讨论了早期贝卡利亚和边沁的观念，see Steven Shavell, *A Note on Marginal Deterrence*, 12 *International Review of Law and Economics* 345 (1992). 关于“边际威慑”的正式论述，see David D. Friedman & William Sjostrom, *Hanged for a Sheep: The Economics of Marginal Deterrence*, 22 *Journal of Legal Studies* 345 - 366 (1993).

〔30〕 Richard A. Posner, *Killing or Wounding to Protect a Property Interest*, 14 *The Journal of Law and Economics* 201 - 232 (1971).

先毫无征兆。正是由于即兴杀人的风险最不可控,所以才会给人们造成最大的恐惧。药家鑫案之所以引来一片喊杀之声,未被媒体和学界阐释清楚的原因就在于此。^[31]

三、演化分析

从要件识别到变量评估,之所以被认为是刑事司法应对要件失灵、破解定性难题的必由之路,是因为诸如破案率和作案成功率之类的变量属于刑法的内生变量,在刑法制定之初就已渗透进刑法条文、教义和犯罪构成之中。但较真的法律人仍可能提出质疑说没有任何证据表明立法者在制定法律时使用了经济分析的方法。确实,从有限的资料记载中,人们更容易发现立法者热衷于谈论的,是那些诸如“良知”“刑法公正”或“校正正义”之类的概念,而不是破案率、作案成功率以及作案可能性等决定刑罚威慑水平或犯罪预期损失的变量。

虽然回应这个质疑很容易,简单粗暴的方式是只需表明,法律经济学的解释力是独立于法律决策者的清醒意识的,^[32]例证俯拾皆是。甚至可以加上一个类比说:星体不懂物理学,并不妨碍物理学解释星体的运动。但要认真回应这个质疑,就必须讲清楚法律和道德直觉以及法学理论之间的历史关联,而这意味着下文将尝试的论证工作具有知识考古的性质,并需要借助一点进化论。

(一) 法律与道德直觉

法学是个古老的学科,但法律的历史却比法学更古老。强调“法律先于法学”的事实,并不是要否认法学对于指引或优化法律决策所有拥有的功能,而只是表明,对于完成法律决策(包括立法、司法和执法)而言,专业化的法学知识不是必需的。但随之而来的问题却是:在缺乏专业化法学知识的条件下,法律决策者如何保证其决策质量?问题的另一面是:仅靠专业化的法学知识,立法者就能创造出伟大的法典吗?

人类最终拥有如此复杂精致的法律制度简直是个奇迹。卢梭很早就意识到,法律的复杂性超出了人类理智所能达到的极限,他曾感慨说:“要为人制定法律,简直需要神明。”^[33]二十世纪中叶,哈耶克从生物学那里获得了启发——既然像生命有机体和生态系统这等复杂精致的结构可以独立于任何强大的心智,法律的起源与演变未尝不可以借助进化的力量。^[34]而只要进化的力量介入其中,只需时间足够长,历经无数次试错、检验和淘汰,简单粗糙的法律制度就能逐渐趋于复杂完备。当然不能无视人类心智在其中发挥的作用,但从结果上看,心智的功能也只是加快了制度进化的速度或缩短了制度进化所需要的时间。

制定法律通常会有一大堆人合作贡献他们的智力,但绝大多数法律都有自己的前身,因此立法其实不那么费事——无须从头做起,并非平地起高楼,只是旧房翻修、添砖加瓦而已。“法系”的概念描述了不同法律的亲缘关系,其间的遗传和变异都清晰可辨。考古学家提供的最早成文法样本是《汉穆拉比法典》,它是3700多年前巴比伦国王汉穆拉比在位时颁布的,因被刻在一个两米多高的石柱上而得以幸存至今。虽然这部法典早已死亡,但它在保护人身和财产、维护社会秩序以

[31] 关于药家鑫案的一个完全不同的解释,参见苏力:《是非与曲直——个案中的法理》,北京大学出版社2019年版,第97—134页。

[32] 经济学的解释力独立于经济活动参与者的清醒意识,见前注[18],波斯纳书,第32—33页。

[33] [法]卢梭:《社会契约论》,何兆武等译,商务印书馆1982年版,第50页。

[34] 参见[英]哈耶克:《致命的自负》,冯克利译,中国社会科学文献出版社2000年版,第73—100页。

及处理纠纷等许多方面体现出的法律智慧却至今存活，并已扩散到世界各地。制度基因可以像生命基因一样永垂不朽，尽管它们的载体已经或即将前仆后继地走向死亡。

很早就有人注意到文化的遗传现象，还专门给文化的遗传因子取了一个名称叫“模因”（meme），对应但有别于“基因”（gene）。^{〔35〕}但法律制度不同于一般的“文化”，其遗传因子和生命有机体的遗传因子很难区分，反而不如混为一谈，干脆就叫作“制度基因”的好。举个例子说，法典中的许多条文体现了“以牙还牙”（Tit for Tat，下文简称“TFT 法则”）的逻辑，但若往前追溯，作为抑制种群内部冲突、促进共生关系的一个制度性的遗传因子，^{〔36〕}TFT 法则早在亿万年前就已经广泛扩散于陆地、海洋和天空的生物圈，或者说，只有那些操控生命体的行为恰好遵循了 TFT 法则的基因，才能在种群中成功扩散并繁衍至今。^{〔37〕}

制度基因的历史比法律更古老，法律的历史比人类更古老。看看我们的近亲黑猩猩或倭黑猩猩的群居生活，就能大概知道，在人类尚未成为人类之前，就已经生活在一个有秩序的群体之中了。^{〔38〕}自然选择留给我们的祖先一笔丰厚的制度遗产，文明的进程因此不必从零起步。那些作为文明支柱的社会规则，非但不一定被表达成文字，甚至无须人类有意识地遵守。

人类的心智也是进化的产物，对于塑造大脑的认知结构，自然选择肯定发挥了基础性的作用。^{〔39〕}只要 TFT 法则对于促进合作和抑制冲突的功能不可替代，那么这个制度因子就不难从基因编程进入大脑认知，而这意味着 TFT 法则可以从动物本能演变为人类的道德直觉（不管是有意识的还是无意识的）。而在缺乏专业化法律知识条件下，只要立法者依靠其道德直觉制定法律，TFT 法则就顺理成章地进入了法律条文。《汉穆拉比法典》就是一个贯彻 TFT 法则的成文法样本。^{〔40〕}

立法者很可能不清楚 TFT 法则的功能和制度逻辑，但只要存在进化的压力，且有足够长的时间，理解就不是正确选择的前提。^{〔41〕}蚂蚁不懂积分学，却能根据太阳移动的轨迹和角速度找到回家的直线；亚马逊河的射水鱼不懂抛物线运动，更不懂光线在水中折射的物理学原理，但仍能将喷水准确射中空中的飞蛾。同样，TFT 法则被真正解码不过三十多年的时间，^{〔42〕}至今仍被广泛误解，^{〔43〕}但几千年来人类的法律制度却似乎从未在这个问题上犯下致命的错误——也许曾经犯过错误，在漫长的人类制度文明史上，不排除曾有个别立法者拒绝接受 TFT 法则，但进化的压力会

〔35〕 参见[英] 理查德·道金斯：《自私的基因》，卢允中译，中信出版社 2012 年版，第 213—228 页。

〔36〕 可以理解为一组信息编码，一组指令的集合。既可以存储于基因体现为生命体无意识的本能，也可以存储于大脑体现为生命体有意识的思考。

〔37〕 参见[美] 罗伯特·赖特：《非零和时代》，于华译，中信出版社 2014 年版，第 277—292 页。

〔38〕 参见[美] 弗朗斯·德瓦尔：《黑猩猩的政治》，杨苡里译，上海译文出版社 2014 年版。

〔39〕 这与哈耶克的信条——不是人类理性创造了社会秩序，而是社会秩序创造了人类的理性——颇有相通之初。见前注〔34〕，哈耶克书，第 15—20 页。

〔40〕 且看《汉穆拉比法典》的几个条文。第 196 条：挖去别人眼睛的人也要被挖出眼睛；第 197 条：打断别人骨头的人也要被打断骨头；第 199 条：挖出奴隶眼睛或是打断奴隶骨头的人要赔偿奴隶价格的一半；第 200 条：打掉同等地位者牙齿的人将会被敲掉牙齿。

〔41〕 人类对于制度性知识只能做到“知其然而不知其所以然”，这个观点是哈耶克的立论基础之一。参见前注〔34〕，哈耶克书，第 19—22 页。但实际上，在哈耶克完成其著述不久，法律与经济学运动兴起，二三十年间就基本揭示了全部法律制度的底层逻辑。

〔42〕 1984 年，罗伯特·阿克赛尔罗德解释了为什么合作与秩序可以发生在一个没有公共权力的群体之中，这是对霍布斯提出的“无公共权力则无秩序”的正式的理论回应，尽管在此之前生物学家、人类学家和社会学家都会霍布斯的见解提出了反例。See Robert Axelrod, *The Evolution of Cooperation*, Basic Books, Inc., 1984.

〔43〕 例如，电影《甘地传》的海报上有一句话：An eye for an eye only ends up the whole world blind.

让他们制定的法律成为失败的变种,而失败的变种最终会消失得默默无闻。

TFT 法则的制度生命力可能超出人们的想象,虽然现代刑法早已废止了同态复仇,但它演化出的各种制度变体却存活至今,且相当繁盛。合同法中的相互制约、担保法中的抵押和留置、刑法中的正当防卫和罪刑相适应,以及其他法律允许的各种形式的自救行为等等,细看都能辨识出 TFT 法则的影子。^[44] 虽然国家垄断了暴力的使用权,但检举、告诉以及作证等许多法律制度的设计,仍然会利用受害人的“报仇心切”甚至“睚眦必报”。^[45] 在科幻小说《三体》里,不计代价的复仇天性被作者描述为人类抵御和威慑强敌入侵的主要心理驱动力。半个多世纪以来,也正是这种野蛮的心理驱动力支撑了核威慑的恐怖平衡。

在博弈论中,TFT 策略在既定的囚徒困境中被定义为:“玩家(player)在第一个回合选择合作,而在后续的每一个回合选择对手在上一个回合选择的策略”。如果对手在上一个回合选择合作,TFT 玩家则以合作奖赏合作;如果对手在上一个回合选择背叛,TFT 玩家则以背叛惩罚背叛。TFT 策略的最大功能,就是能迫使对手与自己合作,如果两个玩家都执行 TFT 策略,就会自始至终互利共赢。^[46] 虽然 TFT 策略素以简单透明著称,但若仔细分析,会发现其制度内涵相当丰富。TFT 玩家可以是诸如基因、细胞、病毒、蜂群、公司、国家之类的任何生命体或生命系统,但下文我们只讨论玩家是个体人类的情形,这意味着玩家拥有了意识、情感以及某种程度的记忆和计算能力。

只要 TFT 策略在群体中扩散到一定范围,就会形成鼓励合作、抑制背叛的 TFT 法则,它可以指引或规范个体的行为。合作应该受到奖赏,对手应以合作奖赏合作;背叛应该受到惩罚,对手应以背叛惩罚背叛。奖赏与惩罚会分别激发正面和负面的情感(喜悦和愤怒),情感和理智结合在一起就形成了道德评价。于是合作、奖赏合作以及惩罚背叛都演变成为一种义务,与之相应的背叛以及拒绝奖赏和拒绝惩罚都算是违反了义务。此外,由于背叛既是一种违法行为,也是一种惩罚手段,所以虽然 TFT 法则禁止背叛,但若对手背叛在先,作为惩罚手段的背叛则是合法的。对应于本文的主题,下文只讨论背叛与惩罚的问题。

(1) TFT 法则会区分背叛的程度,虽然其标准尺度是背叛给对手造成的预期损失,但在最简单的模型中,背叛的程度可以用背叛次数来衡量——一次背叛与两次背叛、多次背叛属于不同程度的背叛。按照 TFT 法则,背叛的次数与惩罚的次数相等。TFT 策略的含义是“以一牙还一牙”,而不是“以两牙还一牙”或“以一牙还两牙”。换言之,无论是“以眼还牙”还是“以毛还牙”,都违背了 TFT 法则。由此,我们看到了罪刑相适应原则的雏形。惩罚不是无成本的,更严厉的惩罚成本也更高,因此需要更强烈的惩罚动机;愤怒可以强化惩罚动机,这意味着背叛的程度越高,对手的愤怒也要更强烈。因此在进化过程中,那些能让自己的愤怒程度与对手背叛程度相适应的玩家,更可能赢得生态竞争的优势。

(2) 在信息不对称的条件下,假定背叛行为被对手识破的概率是 50%,那么 TFT 法则就只能

[44] 关于 TFT 法则贯穿法律制度的经济分析,see Donald Wittman, *Punishment as Retribution*, 4 Theory and Decision 209-237(1974). 威特曼认为当代刑法的许多特点都反映了 TFT 法则。

[45] 参见[美]波斯纳:《正义/司法的经济学》,苏力译,中国政法大学出版社 2002 年版,第 220 页。

[46] 在阿克塞尔罗德设计的著名的计算机博弈模拟竞赛中,“以牙还牙”策略大获全胜——它其实没有赢下任何一场竞赛(至多和对手得分一样高),但却赢了全局。计算机竞赛模拟的是一场生态竞争,“以牙还牙”在各种博弈策略中脱颖而出,意味着执行这一策略的生命体会获得自然选择的青睐。只要“以牙还牙”策略能够最终覆盖整个种群,那么任何两个生命体相互合作的局面就会出现在一个没有公共权力的环境中。See Robert Axelrod, *supra* note [42].

演化为“以两牙还一牙”，因为有多少背叛被识破了，就有同样多的背叛被隐藏了。^{〔47〕} 由于更难以识破的背叛必然会激起对手更强烈的愤怒，所以识破概率会影响我们对背叛行为的道德评价——如果难以被识破的背叛是狡猾的，那么容易被识破的背叛就显得光明磊落。将识破概率与惩罚的严厉程度关联起来，就可以为背叛者创造一个类似边际威慑的激励——如果背叛是既定的，那就应该鼓励对手提高背叛的透明度。但这同时意味着惩罚也必须是光明磊落的，而 TFT 法则的优势之一就恰恰在于，作为一个惩罚方案，它是简单透明的，释放出的激励信号清晰而稳定。^{〔48〕} 由此，我们看到了形式正义的雏形。

(3) 在重复博弈的条件下，一组玩家会相互统计对手背叛的次数和频率，以便准确预测对手背叛的概率。惯于合作但从不过或只偶尔背叛的玩家，会被对手标记为“好玩家”；而惯于背叛但从不过或只偶尔合作的玩家，则会被对手标记为“坏玩家”。当用“好坏”评价一个玩家的时候，意味着从对玩家行为的描述转向了对其心智的描述，这同时也是从事实判断转化为了价值判断。反过来看，价值判断就是对事实判断的模糊处理，描述心智就是对描述行为的模糊处理，价值判断和心智描述最终都可还原为对行为和事实的统计形态。好玩家偶尔背叛之所以很容易被宽恕，而坏玩家偶尔合作也依然要被谴责，其原因就是两者都属例外情形，不至于大幅度改变（而只会微调）对该玩家的行为的统计结论以及对其未来行为的预测。^{〔49〕} 说到底，关心过去还是为了面向未来。

(4) 假定背叛表现为一个过程，行为自始至终需要经历一段时间，那么一组玩家就会相互观察对手的背叛过程。如果过程中的行为与背叛的结果相左，背叛就更可能被认定为“无心”（过失）；如果过程中的行为与背叛的结果吻合，背叛就更可能被认定为“有意”（故意）。区分故意和过失的实质性理由同样是统计学意义上的，在过失的情形下，背叛结果的出现是偶然；而在故意的情形下，背叛结果的出现是常态。同样是背叛，为什么故意要比过失更需要严惩？答案同样是为了面对未来——与严惩过失相比，严惩故意可以阻止更多的背叛。^{〔50〕}

(5) 假定背叛依赖于一个行为方案，执行背叛的方案不见得出现背叛的结果，因而背叛成功只是概率性的；并且如果不同的背叛方案会导致不同的背叛成功率，那么对手就会对不同的背叛方案区别对待——那些成功率更高的背叛方案会受到更严厉的惩罚。而对于那些背叛成功率很高的行为方案，在背叛结果尚未出现时，对手就可能要先下手为强。由此我们看到“危险犯”的雏形。

上述对于 TFT 法则的分析，实际上描述了与刑法相关的人类道德直觉的起源和演化，同时也是对道德直觉解码的过程。解码这些道德直觉，无须校正正义或自然法之类的概念，甚至无须故

〔47〕 从惩罚的角度这是合乎逻辑的，因为惩罚的严厉程度必须与识破概率呈反比例关系才能让威慑效果保持稳定；从背叛的角度也同样合乎逻辑，如果识破概率是 50%，那么识破一次背叛就意味着还有一次背叛被隐藏了。究竟是背叛者太狡猾还是惩罚者太愚蠢？这是分不清楚也没必要分清楚的问题。

〔48〕 在阿克塞尔罗德设计的计算机模拟竞赛中，Downing 策略远比 TFT 更加精于算计且深谋远虑，它通过评估对手合作和背叛的概率来决定自己的选择，在每个回合都力图最大化自己的长期收益，但 Downing 却并不比 TFT 的表现更出色。对此，阿克塞尔罗德评价说：“一个人复杂到不可理解是很危险的。”See Robert Axelrod, *supra* note 〔42〕, at 34–35.

〔49〕 参见〔美〕霍姆斯：《普通法》，冉昊等译，中国政法大学出版社 2006 年版，第 45—57 页。

〔50〕 霍姆斯对故意犯罪与提高惩罚严厉性的关系做出了开创性讨论，尤其集中于犯罪故意与犯罪可能性之间的关系，see Oliver Holmes Wendell, Jr., *The Common Law*, in Mark DeWolfe Howe ed., Little, Brown and Company, 1963, p. 52–62. 与本文类似的一个博弈分析，Ariel Rubinstein, *An Optimal Conviction Policy for Offenses that May Have Been Committed by Accident*, in Steven J. Brams & G. Schwödiauer eds., *Applied Game Theory*, 1979, p. 406–413.

意、过失以及动机好坏之类的概念,而只需背叛可能性、背叛成功率、背叛频率、识破概率、背叛程度等衡量背叛预期损失的变量,道德评价的绝对尺度就是背叛的预期损失;相反,在无力解码道德直觉的时候,反而需要借助校正正义或自然法之类的概念去描述或指称道德直觉。^[51]

至于人类的道德直觉有多少成分属于基因编程,又有多少成分属于大脑认知,这个问题并不重要,重要的是道德直觉本身的功能和逻辑。基因和大脑都是强大的数据处理器,而人类的道德直觉,说到底就是一个算法,这个算法致力于用合理的惩罚去阻止未来的背叛。那些在进化过程中恰好捕获到了最优算法的个体和群体,争取到了生态竞争的优势——这与那些操控蚂蚁或射水鱼行为的基因如果恰好吻合了物理定律和数学逻辑就能有效扩散是一样的道理。算法的合理性与执行算法的生命体(包括基因、细胞、生物有机体乃至整个社会)是否有清醒意识和计算能力没有关系。

在缺乏专业化法学知识的条件下,立法者只能依靠道德直觉去制定法律。但即使仅仅依靠道德直觉,立法者也能制定出一部比较完备的“复仇法典”(参看《汉穆拉比法典》),复仇法典的基础性逻辑就是 TFT 法则。霍姆斯很早就说法律起源于复仇,^[52]这的确是个相当敏锐的洞察。

(二) 法律与法学知识

从博弈论语境切换到法律语境,合作就可以延伸到所有合法行为(除非合作双方会损害第三方的利益),背叛就可以扩展到所有违法行为(除非对手背叛在先)。刑法起源于国家力量介入法律决策之后,虽有国家意志渗透其中,但还不至于摆脱原始复仇法典的色彩,毕竟人们(包括法律决策者)的道德直觉变化不大,更何况作为算法的道德直觉也远未失灵。^[53] 主流的道德直觉对于法律决策构成了强大的约束,在没有明显收益的时候,法律不能冒然冲撞人们的道德直觉。

除了道德直觉,法律决策者还能进行理性思考——这是模拟后果来做出选择的一种计算能力。但如前文分析所表明的,道德直觉和理性思考的算法是基本一致的,对于做出某个特定的决策,道德直觉更多通过自然选择来划掉失败的选项,而理性思考则更多通过模拟后果(最终也是模拟自然选择)来划掉失败的选项,两者的区别主要是个时间问题。^[54] 人类拥有理性思考的能力之后,自然选择的速度就加快了,但还不至于替代自然选择,毕竟人类的思考能力是有限的。只要发现道德直觉、理性思考和自然选择的算法基本一致,我们就只需关心法律决策的功能和逻辑,而无须考虑究竟是哪种力量、起了多大的作用。

刑法的发展趋势,是犯罪被分得越来越细,区别对待的依据越来越多。省略历史细节之后,我们可以粗略想象一下刑法的发展过程。为了制定合理的惩罚方案,法律决策者首先会根据行为和后果区分不同的犯罪。比如,“偷”有别于“骗”,更不同于“抢”。同样是抢,但“抢劫”应该区别于“抢夺”,“侵占”也应从“盗窃”中分离出来。随着犯罪被越分越细,缝隙和重叠就会层出不穷,于是需要设立打补丁的犯罪,比如“寻衅滋事”。除了区分行为和后果,法律决策者还要考虑罪犯的主观心理状态,需要分辨“故意”与“过失”、“蓄谋”与“冲动”,还要区分有无“认识错误”、是否“真心忏悔”等等。此外,立法者还要斟酌罪犯和受害人的状况。比如儿童、精神病人、残疾人和有特定身份的人犯罪,就不能和身心健全的普通成年人犯罪一概而论;同样,如果强奸罪的受害人是幼女或

[51] 前文的分析实际上是借助博弈论的分析工具重述并矫正了十六到十八世纪欧洲论坛上社会契约论关于自然状态中社会规则的分析和描述。

[52] 见前注[49],[美]霍姆斯书,第2—4页。

[53] 关于矫正正义的大量哲学讨论就是重要的验证。例如,John Rawls, *Two Concepts of Rules*, 64 *The Philosophical Review* 3-5(1955); K. G. Armstrong, *The Retributivist Hits Back*, 70 *Mind* 471(1961)。

[54] 参见[英]理查德·道金斯:《盲眼钟表匠》,王道还译,中信出版社2012年版,第45—78页。

盗窃罪的受害人是金融机构，就不能与一般的强奸和盗窃等量齐观。^{〔55〕}

上述区分会让我们不由自主地想象出一张表格，表格的横向是各种类别的罪名——诸如杀人、放火、强奸、抢劫、盗窃、侵占之类；表格的纵向是区分不同犯罪的依据——包括行为主体、危害行为、犯罪对象、危害结果、因果关系、阻却事由以及期待可能性等。传统刑法学理论主要在表格的纵向发挥想象力。把这些区分依据模块化，就出现了“四要件”或“三层次”之类的犯罪构成理论。但这张表格仍然是“犯罪构成理论”基本框架，传统刑法学的主要工作就是把表格填满。

请注意，刑法先于刑法学，世上先有刑法，后有刑法学。面对一部现成的刑法，理论家有两条研究路径。一条是揭示刑法的底层逻辑，将所有犯罪的分类依据还原为决定犯罪预期损失的各种变量，至少要解码人类的道德直觉，毕竟这是最初的立法依据。但由于这条路径难度很大，尤其在“演化博弈”的观念诞生之前肯定是走不通的，而摆在眼前的一条捷径却充满了诱惑，那就是通过填充犯罪分类表来描述不同犯罪的行为、后果、目的、动机，犯罪意图、主体、客体、人身危险性以及受害人等各个方面的状况。从一种犯罪中发现的共同点被整合在一起，就变身“刑法分论”中不同罪名的构成要件；从所有犯罪中发现的共同点被整合在一起，就变身“刑法总论”中犯罪构成理论。

那么，如何评价传统刑法学理论的质量呢？首先要承认它的功能，对于指导定罪量刑，传统刑法学理论至少在统计学意义上具有合理性——如果过去发生的某一类犯罪都体现出某些共同点，那么未来发生的同类犯罪体现出这些共同点的概率就很高，因而，将体现出这些共同点的未来犯罪认定为同类犯罪，就通常不会犯错误。举例而言，如果过去发生的盗窃罪都是以隐蔽手段侵占别人的钱财，那么未来发生以同样手段侵占别人财产的行为被定性为盗窃，就通常不会犯错误。生物学也有同样的逻辑，如果已知的所有昆虫都长着两对翅和三对足，那么如果未来发现某个未知动物也有同样的形态，就大致可以断定这种动物是昆虫。^{〔56〕}

但概率性的推测总会有例外。以隐蔽手段侵占他人财产未必就是盗窃（比如受害人是自己亲属的情形），盗窃也未必使用隐蔽手段（尤其在监控摄像普及之后）。同样，长着两对翅和三对足的动物未必是昆虫（尽管目前还没发现），昆虫也未必都有两对翅和三对足（有些昆虫就是无翅的或只有一对翅）。但只要诸如此类的例外数量稀少，就不至于破坏大局。但两相比较的区别却显而易见，生物学家从来不会把两对翅和三对足视为昆虫的“构成要件”，但法学家却会把“隐蔽手段”和“非法占有”视为盗窃罪的构成要件。^{〔57〕}

传统刑法学之所以无力应对要件失灵，就是因为要件失灵属于例外情形。回避被道德直觉、理性思考和自然选择共同支配的刑法底层逻辑，是这种理论的先天不足。无论刑法总论中的犯罪构成，还是刑法分论中的构成要件，都只是以一种简单、生硬、随机的方式组合在一起的。貌似一种完整的理论，细看则是残垣断壁。而与此同时，决定这些要件的变量以及不同变量之间的函数关系却被掩盖了。

这种情形不仅导致刑法概念体系的混乱不清，而且割裂了彼此之间的关联。比如：“犯罪行为”和“犯罪情节”如何区分？“危害后果”和“社会危害性”是怎样一种关系？“人身危险性”和“社

〔55〕 区分受害人，是因为某些犯罪受害人（比如幼女或金融机构）更容易受到侵犯，或一旦受到侵犯其损失就更大，归根到底还是受害人不同会导致犯罪的预期损失。

〔56〕 可以通过外表形态的典型特征去识别昆虫，但不能断定诸如“三对足、两对翅”之类的特征为昆虫专有，考虑到世界上还有大量未知物种，尤其是浮游生物，所以只能说“三对足、两对翅”对于识别昆虫具有统计学意义上的合理性。

〔57〕 这种思维方式曾经导致错误，例如 2013 年关于行为人不知道幼女实际年龄而发生性关系不以强奸论的司法解释，参见苏力：《司法解释、公共政策和最高法院——从最高法院有关“奸淫幼女”的司法解释切入》，载《法学》2003 年第 8 期，第 3—29 页。

会危害性”是两个相互独立的衡量标准吗？如果不是，两者之间如何换算或如何解决重叠的部分？“违法性”和“该当性”又是怎么回事？如何解决逻辑上的循环论证？“非法占有”可以作为一个要件吗？“故意”和“过失”之间界限分明吗？如果不是，两者之间如何过渡？“认识错误”属于过失的一种，与其他过失又有什么区别？“主观方面”和“客观方面”有因果关系吗？如果有，那么“主客观相统一”岂不是把因果关系给掩盖了？^{〔58〕}……诸如此类的问题实在太多了。如今刑法学（其实是整个传统法学）的概念体系像极了牛顿之前的物理学（那时还叫“自然哲学”）的概念体系，其中充满了歧义和混乱，而且难以度量。

对比一下“生物分类学”，传统刑法学理论的先天不足就暴露得更清楚了。最早的生物分类学是按形态分类，比如中国古人就把动物分为虫、鱼、鸟、兽四个大类。^{〔59〕}想必达尔文最初看到三对足的无翅虫时，也曾犯过犹豫，这些虫子更接近于蜘蛛还是更接近于蝴蝶？类似的问题是，海豚更接近于鲨鱼还是更接近于猎豹？蚂蚁更接近于白蚁还是更接近于蜜蜂？生物分类学通过研究生物类群间的异同和异同程度来阐明生物间的亲缘关系、进化过程和发展规律，但要判断两种生物之间的亲缘关系，只看形态是靠不住的。生物分类学从古老的形态学发展到现代的分子生物学，其分类的依据也从外表形态转向更微观的标准，比如染色体和血清反应等。^{〔60〕}于是我们知道了，海豚和猎豹的亲缘关系更近，虽然它更像鲨鱼；而蚂蚁和蜜蜂的亲缘关系更近，虽然它更像白蚁。三对足的无翅虫与蝴蝶而非与蜘蛛或蜈蚣的亲缘关系更近，因此被列入昆虫之中。

无论是以染色体为依据的细胞分类学，还是以血清反应为依据的血清分类学，相对于古老的形态学，都是“微观理论”（microscopic theory）和“现象理论”（phenomenological theory）之间的关系。科学研究经常从“现象理论”起步，然后被“微观理论”取而代之。相对于前者，后者有能力揭示更深层的因果关系，具有更强的解释力、精确性和整合能力，甚至可能更简洁。牛顿的经典力学相对于亚里士多德甚或伽利略的力学理论，就是微观理论和现象理论之间的关系，而当量子力学出现之后，经典力学就变成了一种现象理论。

传统法学和法律经济学同样是现象理论和微观理论的关系。在传统刑法学的框架中，犯罪的行为、后果、主观心理状态、主体、客体等等都是分裂的，而法律经济学却把这些概念还原为决定犯罪预期损失和刑罚威慑水平的各种变量，从而将这些分裂的概念整合在一起。^{〔61〕}比如过失和故意、蓄谋和冲动，都是分裂的概念，但经济分析却可以把它们整合进事故预期损失和事故预防成本的函数关系中，并通过作案可能性、作案成功率、破案率以及犯罪实际损失等变量去表征不同程度的犯罪预期损失。^{〔62〕}从要件识别到变量评估，类似于生物分类学从形态识别转向比对染色体或观测血清反应，倘若没有这种从现象理论到微观理论的转向，我们不可能发现强吃霸王餐的行为

〔58〕 牛顿将一些古老但却意义模糊的概念（比如“力”“运动”，甚至“时间”）赋予新的含义，由此开启的物理学的新时代。这些概念变得可以量化，因此可以被用于数学方程。而在此之前，“运动”（motion）一词可以指代诸如“桃子成熟”“石头落地”“孩子成长”以及“尸体腐烂”等极其广泛的现象。参见〔美〕詹姆斯·格雷克：《信息简史》，高博译，人民邮电出版社2013年版，第5页。“能量”（energy）和“信息”（information）等科学概念也经历了类似的变化。

〔59〕 汉初《尔雅》就是这种分类，有“释虫”“释鱼”“释鸟”“释兽”等条目。

〔60〕 比对染色体为依据，产生了细胞分类学；以观测血清反应为依据，产生了血清分类学。相对于原始的形态分类学，都算是以微观理论取代现象理论的进展。

〔61〕 Stephen G. Gilles, *On Determining Negligence: Hand Formula Balancing, the Reasonable Person Standard, and the Jury*, 54 Vanderbilt Law Review 813–861 (2001).

〔62〕 关于故意和过失，see Steven Shavell, *Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent*, 85 Columbia Law Review 1232–1262 (1985).

其实更接近于抢夺，尽管它看上去更像抢劫。

当然，传统刑法学理论并非一无是处，说它要被法律经济学取而代之也是言过其实。正如我们看到的，传统刑法学理论可以顺利解决大部分格式化的案件，那些典型的盗窃、抢劫、诈骗，根本无须变量评估，只需简单比对一下要件就足够了。就像我们教会儿童怎样识别那些典型的昆虫，比如蝴蝶、蝗虫或苍蝇，通常只需数数翅膀和腿足就够了，用不着去比对染色体或者观测血清反应。

但“顺利解决”不等于“高质量解决”。不可否认，即使要件完全相同的两个案件，细节也是千差万别的，但在传统刑法学的理论视野中，要件之外的细节却仿佛被打上了马赛克。要解读马赛克之内的细节差异，法律决策者只能重新依靠他们的道德直觉。许霆案之所以引起轰动，不是因为法律指向的判决结果违背了这种犯罪的构成要件，而是因为判决结果严重冲撞了人们（包括法学专家）的道德直觉。

对此，刑法学教义能否会发挥些作用？只能说，作用很有限。对于个案而言，一条教义的指引功能充其量相当于一个法条，并且它还像法条一样对事实进行条块分割。这意味着，即使刑法教义有助于缩小马赛克的区间，也是杯水车薪，毕竟细节是个无穷大。更何况马赛克被缩小的代价是增加了缝隙和重叠。

在马赛克的区间之内法律决策者可以自由裁量，而自由裁量的隐含义就是允许法律决策者根据自己的道德直觉来操作法律。然而见仁见智，法律决策者的道德直觉不仅因人而异，而且很难与时俱进。受制于进化的速度，我们的道德直觉很可能还停留在农业社会，甚至狩猎采集社会，至多是工业社会的老样子，与信息时代社会变迁呈现出的加速度难以完全合拍。

而如前文所述，道德直觉是一个算法，只有把这个算法解析出来，才能从基因操控的下意识层面提升到意识层面，从自在变成自觉，变成评价罪责轻重进而指引法律决策的规范性思考，从而补充、矫正乃至替代法律决策者的道德直觉，进而创造出的一套可以交流的语言操作系统，将见仁见智的道德直觉收敛、聚合到同一个频道。

虽然法律人会说，受认知能力和成本的制约，变量评估的方法不可能做到精确，评估结果也会因人而异。但这个质疑似是而非，或者说，质疑的内容是对的，但却没有质疑的力量。方法和理论的优劣都是相对而言的，世界上没有哪种理论可以让法律决策变得高度精准、不差毫厘，有思路不等于捕捉最优，但有思路总比没思路好，因为没有思路会大大降低捕获最优的概率。且评估优于描述，把需要评估的变量解析出来，比之继续使用掩盖变量的描述性概念可谓前进了一大步。如果说变量评估做不到精确，那么那些充满了混乱和歧义的描述性概念岂不是更加离谱了？更何况，法律决策通常并不需要精确的评估结论，法律决策面对的难题通常只是“某种违法行为更接近于犯罪 A 还是更接近于犯罪 B”或者“处理某个案例应该采用方案 A 还是方案 B”，对应于此，变量评估通常只需分出个大小，因而方案评估通常只需要知道个好歹，就足够了。

四、结 语

对于如何破解司法实践中的定性难题，本文提出了用变量评估取代要件识别的方法来应对要件失灵，并以回应质疑的防御姿态论证了这种方法的合理性，其理论支撑显然是法律经济学。

虽然没有解码人类的道德直觉，但将法律视为致力于财富最大化的社会治理工具，并利用理性选择理论，法律经济学也照样揭示了法律的底层逻辑，两者算是殊途同归。只有法律的底层逻辑显现出来，那些被法条、教义以及要件分割的马赛克才可能彻底消失。法律决策者由此心明眼亮，更高的像素自然会呈现出一片清晰的视野。

多少有点意外的是,本文的讨论还可以为当下的一个学术热点,即实际上已经演变为一场运动的“法律与人工智能”提供一些启发。这场声势浩大的学术运动似乎遗忘了一个关键问题:如何为人工智能操作法律设计出一套算法?

如果考虑到被称为“涵摄”的三段论推理显然不顶用,而号称利用了大数据的统计学算法又难免系统性偏差,那么就会发现,真正值得期待的,是将法律本身算法化,人工智能在掌握这套算法之后才拥有操作法律的能力。即使退而求其次,也至少需要设计出拥有“道德判断力”的人工智能,解码道德直觉的任务因此责无旁贷。但不管怎样,抛开法律经济学或演化博弈论去讨论“法律与人工智能”都是令人费解的。

想象一下,如果将来有一天,强大的人工智能能够替代人类去完成法律决策。只要信息科技和生物科技能把事实探知的所有障碍彻底扫除,人工智能就不必遵循证据法规则,也无须了解程序正义的观念。面对一种犯罪行为,它可以通过评估关键变量计算出罪行轻重,然后根据可描述“罪刑相适应”的一个单调函数直接做出处罚。定罪环节就被彻底省略了,因为定罪的原始目标就是为了准确量刑,因而基于算法的量刑只要足够准确就可以取代定罪。^[63]

这个变化是革命性的。那些识别不同犯罪的要件、区分不同犯罪的依据,以及描述不同犯罪行为、后果和主观意图的一系列概念,甚至包括不同的罪名,都会成为历史词汇。剩下的只有评估罪行轻重的各种变量,比如破案率和作案成功率。但这个革命性的变化不会让法律飞向天际,而只是把法律带回到它的起点——那个最原始的状态。

Abstract The law predates its theories. In criminal law, the society's moral intuition about culpability had served as the original ground on which criminal law was built. The formal theory of criminal liability, however, has mostly bypassed the substantive analysis of the law's underlying moral intuition and has adopted the analysis of the assignment of criminal liability by putting together essentially a checklist of elements for each offense instead. Although such an approach appears practically useful and is also often supplemented by legal dogmatics, it falls short from time to time particularly when solving the hard cases where multiple types of offenses appear plausible for the defendant's act. Only by focusing on the underlying logic of criminal law may one appreciate that the critical rationale for culpability lies in the magnitude of a crime's expected loss. Therefore, all decisions on criminal conviction and sentencing must ultimately require an examination of the detection rate, the crime success rate and other variables that affect the level of deterrence and the expected loss from a crime. Replacing the element checklist with a methodology that focuses on analyzing these substantive variables may effectively solve the practical difficulty of differentiating offenses for criminal conviction.

Keywords Formal Theory of Criminal Liability, Elements of Crime, Robbery, Fraud, Marginal Deterrence

(责任编辑:宾凯)

[63] 这和柏拉图在《理想国》中对“哲学王”的想象颇有想通之处。在《法律篇》中,柏拉图之所以退而求其次地接受了法治,就是因为意识到社会不具备“哲学王”的信息条件。