

# HOW NUMBERS RULE THE WORLD

## THE USE AND ABUSE OF STATISTICS IN GLOBAL POLITICS

我们不应轻信数据永远能揭露事实 数据无法拯救我们，我们只能依靠自己。

# 大数据战争

## 数据在全球市场的使用与滥用

【英】洛伦佐·费尔拉蒙蒂（Lorenzo Fioramonti）/著 张梦溪 译

**数据并非  
动态的实体**

相反，它们的本质是静止的

**数据带来  
“博学的无知”**

让市场成为“概念监狱”

数据的微妙之处在于  
它们并不消灭权力，  
而是将它隐藏起来

**数据抹杀了我们  
批判性思维的  
集体能力，助长  
愚蠢的风险**

How Numbers Rule  
the World  
The Use and Abuse of Statistics in Global Politics

大数据战争  
数据在全球市场的使用与滥用

【英】洛伦佐·费尔拉蒙蒂 (Lorenzo Fioramonti) / 著  
张梦溪 译

中华工商联合出版社

图书在版编目（**CIP**）数据

大数据战争：数据在全球市场的使用与滥用/（英）洛伦佐·费尔拉蒙蒂著；张梦溪译．—北京：中华工商联合出版社，2018.7

书名原文：HOW NUMBERS RULE THE WORLD: THE USE AND ABUSE OF STATISTICS IN GLOBAL POLITICS

ISBN 978-7-5158-2354-6

I.①大... II.①洛...②张... III.①数据处理 IV.①TP274

中国版本图书馆CIP数据核字（2018）第128161号

Original title: How Numbers Rule the World: The Use and Abuse of Statistics in Global Politics

Copyright © Lorenzo Fioramonti 2014

Published by Zed books in 2014

All rights reserved.

The simplified Chinese translation rights arranged through Rightol Media（本书中文简体版权经由锐拓传媒取得 Email: [copyright@rightol.com](mailto:copyright@rightol.com)）

北京市版权局著作权合同登记号：图字**01-2018-4655**号

大数据战争：数据在全球市场的使用与滥用

How Numbers Rule the World: The Use and Abuse of Statistics in Global Politics

---

作 者：【英】洛伦佐·费尔拉蒙蒂（Lorenzo Fioramonti）

译者：张梦溪

责任编辑：于建廷 臧赞杰

责任审读：傅德华

营销总监：姜 越 郑 奕

营销企划：张 朋 徐 涛

版权推广：袁一鸣 吴建新

装帧设计：周 源

责任印制：迈致红

出版：中华工商联合出版社有限责任公司

发行：中华工商联合出版社有限责任公司

印刷：三河市宏盛印务有限公司

版次：2018年9月第1版

印次：2018年9月第1次印刷

开本：700mm×1020mm 1/16

字数：230千字

印张：14.75

书号：ISBN 978-7-5158-2354-6

---

服务热线：010-58301130

团购热线：010-58302813

地址邮编：北京市西城区西环广场A座19-20层，100044

**Http:** //www. chgs lcbs. cn

**E-mail:** cicap1202@sina. com（营销中心）

**E-mail:** y9001@163. com（第七编辑室）

工商联版图书

版权所有 盗版必究

# 目 录

---

## [导言](#)

## [第一章 数据的力量](#)

### [数据与政治](#)

### [相信数据](#)

### [数据、市场和民主](#)

## [第二章 新的全球统治者：信用评级成为不可控制的力量](#)

### [信用评级机构：从市场分析师到寡头垄断者](#)

### [数据背后：非正当的交易](#)

### [评级政策和全球金融危机](#)

### [评级与非理性](#)

## [第三章 数据对全球变暖袖手旁观：气候变化的市场化](#)

### [环境怀疑论和成本效益分析的兴起](#)

### [“气候门”事件：对气候数据的扭曲](#)

### [气候市场](#)

### [结论：当数据成为危险的干扰](#)

## [第四章 衡量不可测：数据把自然金融化](#)

### [衡量自然的价值：全球治理的数据化演变](#)

### [生态系统的价格评估](#)

### [自然界的金融化](#)

### [结论：自然有限公司的成立](#)

## [第五章 数据能够为人类带来利好吗？](#)

### [援助效力的政治：历史回顾](#)

### [对于证据的需求](#)

### [慈善资本主义的崛起](#)

### [数据的罢工：社会领域影响评估的政治](#)

### [“行善”市场化](#)

## [结语 重新思考数据，重新思考治理](#)

### [为数据辩护](#)

### [在优劣数据之外](#)

### [治理、数据与公共领域](#)

现在怎么办？

致谢

谨以此书献给我的孩子们

和他们成长的这个世界



---

## 导言

# 你被数据控制了吗？！

---

“我们不断地利用我们不理解的公式、符号和规则，以使我们能够得到那些我们个人所未掌握的知识之帮助。”

——F. A. 哈耶克《知识在社会中的利用》，1945年

“最重要的东西是无法衡量的。”

——爱德华兹·戴明

在浪漫喜剧电影《我配不上她》中，男主角问自己：“10分的人怎么会喜欢5分的人？”电影围绕着一段爱情故事展开，讲述了一个平凡的机场安检员柯克和从律师变为活动策划人的漂亮女孩莫莉之间发生的事。就人物魅力而言，莫莉绝对是一个“10分”女孩，她美丽、聪慧而富有。而男主角是一个低于平均水平的“5分”男人，他十分瘦弱，只是一个工薪阶层的年轻人。虽然柯克爱上了莫莉，莫莉显然也对他很感兴趣，但他们在相处中却表现得十分笨拙，发生了许多意想不到的事情。就像柯克的朋友说的那样，一个10分的人也许能和8分的人交往，5分的人也可能与7分的人接触，但5分的男人绝对配不上10分的女人。

数学本身并不存在引申的概念，但可见的是其中的差距实在太太大。数据的规则就是这么明确。电影中的所有配角，包括柯克和莫莉的家人、好朋友和各自的前任都认为他们俩的恋爱关系不可能维持下去，试图说服两个年轻的恋人放弃彼此。不出所料（但有趣）的是，在接下来的剧情中，两个主角逐渐疏远，互相猜疑。他们之间的交往变得不自然，显而易见的琐碎小事被不断放大，误解也无处不在。两个恋人对此感到困惑。他们不再了解彼此的感受，在关系破裂后重新回到了各自的世界。10分的人和5分的人不可能成为一对幸福的伴侣，至少看起来是这样。当然，在影片的结尾，柯克和莫莉还是重归于好，从此过上了幸福的生活。尽管有一些不太好的事情发生，爱情仍然获得了胜利。两个主角克服了纠结

于分数的毛病，选择听从自己的内心。在数学上不可能发生的事，被爱的不理性击败了。爱情喜剧就是这样，因为可预见的狗血结局而落入俗套。

根据英国统计与生物学家罗纳德·费舍尔（Ronald A. Fisher）的观点，统计学是“人类进步的特别方面”，“将其特殊性赋予了二十世纪”。无论我们是否同意这种说法，我们不能否认数据在很大程度上操纵着我们的世界，成为我们社会、经济和政治决策背后的推动力量。正如柯克和莫莉一样，数据影响着我们和我们身边人的行为。我们每天都会进行衡量和比较，不断地通过通用的尺度来评价自己的美貌、智力、机敏和成功。我们会量化所有事物，包括收入、性能力、生活质量和幸福程度。我们的生活被数据包围着。当我们习惯这一切，就不再注意到数据的力量。

我们来看一下一些研究所和私人机构在环境政策、社会福利和治理等不同领域制定的具体指数和指标。我计算了一下，社会科学领域有300多个合计指数，它们来自于数以万计的单项指标，事实上肯定还不止这些。这还不完全包括政府统计部门的那些数据。而且，这种数据是非常不完整的，它只涵盖了一些特定的研究领域，比如我在自己的学术经历中涉及的社会发展和治理，甚至没有纳入由自然科学家、工程师、医生、建筑师和其他“硬”学科代表所产生的数据，而这些数据可能会有无限多。

每个星期，每个月或者每年，世界各地的几百个智库、非政府组织（NGO）和研究中心都会生成一定量的数据来评估经济状况、城市生活质量、企业竞争力、低收入国家的发展速度、教育水平、福利体系中提供的服务、给社会带来破坏的贪污腐败以及无数社会机构的工作等。他们甚至会测量一些明显是“微不足道”的经济方面，比如全球范围内麦当劳巨无霸的比较成本、不同财富得以跨境的便利程度、公司在某个国家获得执照的所需天数、所有可能的国家风险（包括战争、恐怖主义和经济不稳定性）、银行评级、公司和国家债务、旅游竞争力、计算机素养、受教育程度、全球性的饥饿、粮食安全等。

这些数据用来评估发展战略、衡量绩效，为决策提供信息和指导改革。简言之，它们推动了全球治理和国家治理。比如说，绩效指标已经成了分配对外援助和投资的重要标准，于是它得以在很大程度上影响一国的经济水平。这些数据带来的政治、经济影响力也许在私人机构发布的信用评级中得到了最明显的体现。同样，国际货币基金组织（IMF）发布的财政评估数据能单方面强制各国遵从。

在许多部门，数据的激增已经模糊了公共和私营机构之间的区别，如政府制定的制度、程序，公司、机构、顾问、审计员和非政府组织制定的数据基准、评级和排名往往会对其进行补充（且经常得以实施）。为了获取数据在政治和经济事件中的影响，并结合公共和私营机构的作用，许多分析师引入了“元治理”（Metagovernance）和“跨国私人治理”（Transnational private governance）等术语。

统计测量在如今有着无限的应用。数据在所有领域都发挥着主导地位。即使是体育运动，也在很大程度上依赖着数据的考量，专家会在电视节目中量化运动员的表现和战术模型。棒球就是一个我们熟知的例子，它的传统和数据紧密相连。现代棒球之父亨利·查德威克（Henry Chadwick）在他1868年出版的《棒球运动：如何学，如何打，如何教》一书中，系统地将统计推断应用于棒球规则，而后，棒球运动很快成了全美国最流行的运动。查德威克曾是一名板球专项记者，他热衷于数据统计，并创造了世界上第一个棒球数据集。他列出了所有比赛的具体细节，包括主要俱乐部击球手的出局、得分、全垒打和三振的次数以及频率。他认为，其中一些统计数据能够反映击球手的真正能力，而另一些则不能。他不赞成对全垒打（所有种类的全垒打）过分“炫耀”，并且强烈建议在计算打击率时将场上失误击打排除在外，因为这显然没有表现出他们的击球技巧。他想利用统计数据来革新这项运动，使之“从二十年前的简单场地练习，变为如今更具男性化和科学性的球类运动”。

棒球运动中数据的力量为2003年的畅销书《魔球：逆境中制胜的智慧》提供了叙事线索，这本书后来还被改编成了一部由布拉德·皮特主演的好莱坞电影。该书聚焦于奥克兰运动家棒球队的经营策略，利用场上数据进行比赛统计分析（即以美国棒球研究协会命名的赛伯计量学来进行统计分析），在预算有限的情况下，组建一支实力强劲的队伍。《魔球：逆境中制胜的智慧》中的核心前提是，棒球界的传统认知经常会忽略一些重要的潜在因素，而它们却成了决定比赛输赢结果的关键。通常用来评价球员能力和确定薪水的因素包括盗垒成功、打者打点、打击率等。它们对于评价一支队伍的成功与否则显得有些过时了。相比之下，严谨的统计分析表明，像上垒率和长打率这样的其他因素才是评判成功更好的指标，虽然对观众来说这些数据并不瞩目。这样的观测结果往往和传统棒球认知并不相同，于是奥克兰运动家棒球队能够以一个相对适中的价格招募到在这些“新”指标上表现优异的球员，因为棒球市场常常低估了他们的价值。

哈佛大学哲学教授迈克尔·桑德尔（Michael Sandel）在《金钱不能买什么》一书中，对奥克兰运动家棒球队系统性采用现代统计模型进行了论述。他认为，这支队伍“把新型量化交易者带给华尔街的东西运用到了棒球比赛中，而这就是利用计算机分析战胜依赖天分和个人经验的老将的能力”。桑德尔通过分析奥克兰运动家棒球队的例子，展现了统计推理和市场之间的深切联系，特别是在量化方法被用于产生更有效的定价机制和在对抗中获得竞争优势的情况中。奥克兰运动家棒球队的战略的确一度获得了成功，在2002年，这支队伍获得了美国联盟的西部冠军。但是，当其他队伍亦步亦趋，雇佣最好的统计人员，比财力不足的竞争对手出更高的价，奥克兰运动家棒球队反而因为它的成功成了受害者。如今，像波士顿红袜队和纽约洋基队这种薪金水平最高的球队，也会经常使用计算机统计模型来制定比赛策略，评估球员价值。而比赛本身也发生了一些改变，因为统计数据的功能已经逐步取代了球员比赛时的本能，比赛本身的激情逐渐被消磨。矛盾的是，“在决定大联盟各球队的比赛胜率方面，金钱渐渐变得更重要，而非相反”。

我们能从棒球比赛中得到这样的启发，在我们的社会和政治生活的所有领域，尽管市场公信力受到了来自全球经济危机的打击，数据的激增还是直接或间接导致了市场势不可当的扩张。这反过来又破坏了公共领域的民主参与和审议功能，毕竟它是人们论述、争辩和推广思想的舞台。

这就是本书想说的内容。接下来的章节将展示数据在人们的生活、经济、国家治理过程中如何被使用和滥用，市场的力量被放大，公众的思想被左右。

第一章将通过引入科学史研究，讨论某些数据如何被纳入决策过程。事实上，统计学的科层化是现代国家的一个典型特征，尤其是自19世纪末以来，公共基础设施不断建成，其成本、收费及税务的计量不断增长。在对数据的统治力进行总体概述后，我们将深入探讨第一个实证案例：信用评级机构和它们对全球治理的影响。

第二章将对评级的历史和评级机构在公共财政治理演进过程中产生的利益冲突进行分析。该章通过详细说明信用评级如何被纳入国家与国际政策，发现将评级纳入几乎所有国家的公共治理领域会导致金融市场中某些部分的强化，损害民主的问责制度。

第三章将着手于如何缓和气候变化的棘手问题，这是全球治理的另一个基本领域。在这一领域，统计数据的不同运用引发了气候学家和所谓的“怀疑论者”之间的冲突，每个阵营都用一组数据来佐证不同（若非相反）的事实。两方阵营会对此做不同程度的处理，并最终通过成本效益分析来选择最佳政策，以缓解工业增长对环境的影响。这些分析通过一系列关键假设和计量经济模型，实现成本和效益的货币化，支持了碳市场、补偿方案以及排污权交易的建立。

第四章通过分析自然资本和生态系统服务价值评估的新方法，深入探讨统计数据与环境治理之间的微妙关系。虽然其中一些方法旨在通过将环境退化的代价纳入评价指标来“纠正”国内生产总值（GDP）的计算，它们却也会鼓励金融市场延伸进入自然界，并对世界生态系统造成潜在的危险。

第五章将着眼于某些类型的衡量标准如何影响发展援助部门和通过强化商业部门的宣传工具进行全球反贫困斗争。这种类似于企业的发展方式不仅仅局限于如今的援助机构对发展中国家的作用力和效果的衡量，还影响到了工业化国家在民主社会、慈善基金会非盈利运作模式的影响下对社会变革的推动。

本书并不否认数据对社会进步的重要性。没有统计数据，政策将只由印象主义考量和修辞论证主导。测量是人类生活的基本组成部分。随意列举几个基本发展领域——我们的教育、医疗保健和住房都依赖于测量。但同时，我们不应轻信数据永远能揭露事实。

在社会领域，统计数据总是由批判性假设推动的，这在做出影响整个社会的决定时，应予以考虑。通常，这些假设由狭义计量经济学驱动，并且会以更全面的考虑作为代价。我们可以计算一对夫妇在一起度过的时间、他们谈论的话题、出门的时间、挣的钱、发生性行为的频率等，然而，这些所有的总和并不等同于爱，就如同柯克和莫莉历经艰难认识到的那样。标准化考试在一些学校中可能是有用的，但它不应该被视为对教育教学的全部估量。也许大多数的考试评估能帮我们找出学生的差距，但其隐藏的东西比揭示的更多。这就是为什么一个健康的社会需要区分开能够测量的领域和不能测量的领域。我们能测量的部分总是有限的。如果把数据推理延伸过度，我们会过分简化一些事实。当这种情况发生，我们的测量工具就比测量的内容更重要。

我们最终希望得到的是我们能测量的东西，而非测量我们想要的东西。否

则，学生将只学习与考试相关的东西，夫妻将只关注表达情感的频率而非质量，各种领域的从业者将会被强加于他们的生产力参数所束缚。在其他一些领域，比方说生态治理，各种声称要保护环境的审计员将对自然界进行货币化。而最终，对测量的偏执会导致社会关系和自然世界的商品化。在某种程度上，这就是市场在测量的大数据时代里变得如此强大的原因。作为经济交易的中心，市场更易于测量。它们的概念、原则和功能都非常适合经济和统计分类。

这本书写到了数据在现代治理最关键的领域中如何加强技术统治，数据如何在我们的社会和政治生活中加强市场的控制力。总而言之，书中会介绍数据在减少公众参与和理性辩论中起到的作用，它使我们已经饱受打击的脆弱民主力量变得更加单薄。

---

## 第一章

# 数据的力量

---

“新闻、政治和日常生活中，到处都有数据。无论好坏，它们俨然已成了如今最卓越的公共语言——会说这种话的人，便可统治世界。”

——M. 布拉斯兰&A. 迪尔诺《数据游戏》，2009年

“计数使统计者升级，而使被统计之物降级。”

——罗伯特·钱伯斯《谁的现实有价值？》

古希腊哲学家、数学家毕达哥拉斯出生于萨摩斯岛，他深信数据的力量。他不仅把毕生精力投入了数学学习中（到如今他的定理仍然是几何定理的基本组成部分），还发展了以数据为基础的整个哲学和宗教体系。他在意大利克罗顿建立了“毕达哥拉斯学派”，其信徒在当时颇受欢迎。该学派积极参加政治活动，提出了一些改革，倡导平等的文化观念，包括对妇女和其社会角色非常先进的看法。毕达哥拉斯自己也见证了从哲学家到统治者的角色转换，这一观念后来影响了柏拉图的政治思想，尤其是在他的代表作《理想国》中，他提出了这样一个主张，即“志同道合的思想家结成一个组织严密的公社”，为他们“生活其间的政治组织建立领导，乃至统治”。根据英国哲学家罗素（Bertrand Russell）的观点，毕达哥拉斯将数据视为自然背后的推动力，这深刻地影响了柏拉图哲学和大多数的西方哲学思想。毕达哥拉斯也启发了柏拉图哲学的一个中心思想，“数学与一般抽象思维（包括逻辑）不但可以充当近代意义中的哲学基础，而且可以充当科学与伦理学命题的基础。”毕达哥拉斯和柏拉图都受到了亚里士多德的严厉批判，亚里士多德强调形式（Form）与物质（Matter）之间的重要区别，指出“一项和真理同样重要的谬误是（数据）本身可以建立起有关物质世界的知识”。

虽然毕达哥拉斯生活的细节鲜为人知，但我们知道他在政治目的上对数据的使用主要都基于说服行为。他旨在建立和运行一个完美的社会，在这个社会里，冲突和意见多样性被认为不利于社会稳定和维护秩序。他的政治生涯大约持续了

二十年，最终结束于革命。考虑到当时不稳定的政治制度，这已经是相当长的一段时期。在变革中，毕达哥拉斯学派的活动场所遭到了严重破坏，许多门徒被杀害。毕达哥拉斯本人被迫逃往克罗顿，在流放中身亡。然而，他的遗产对许多秘密组织的传统产生了长久影响，包括中世纪建立的两个秘密组织——共济会和玫瑰十字会。

Validity（效度）的拉丁词根validitas意指“优势与力量”。一个人的效力即是他能够命令别人服从的能力。也就是说，效力就是力量。然而，效度也是统计学的一个基础概念，用以表明一种测量方式能在多大程度上准确符合被测量的事物。在社会研究中，效度的定义有很多种（例如表面效度、效标效度、结构效度等），这表明了测量方式需要尽可能地与它们希望描述的潜在现象相一致，以对现实生活产生有意义的推论。这里的关键假设是统计数据能在我们感受世界和获取知识的过程中产生重要的影响。由此，这些统计数据才得以影响我们的决定，以及我们如何统治作为一个集体的我们自己。于是，效度对于数据的真实性来说至关重要。而与此同时，效度也在说服力上得以体现，就如同在毕达哥拉斯时代展现的那样。

公共统计在维持社会秩序和权力结构方面至关重要。衡量国家财富的初次尝试（可以说是当代国内生产总值的雏形）由经济学家威廉·配第（William Petty）在1652年做出，它成为奥利弗·克伦威尔（Oliver Cromwell）向爱尔兰军队承诺的土地再分配计划的一部分。这项调查旨在为英国政府的利益服务，它的目标是通过征用该国民众（尤其是天主教徒）的生产性土地，并将其变成永久军事占用的一部分收入来源，以解决爱尔兰问题。一些历史学家已经证明，这项统计工作在多大程度上根除了爱尔兰的土著文化，其他一些学者将其描述为“原始积累的庞大实验”。配第的工作也有助于为政府提供新的信息来增加税收，并限制私人拥有的财富，这是一种控制地方自治权的有效智慧，避免了资本集中在潜在的对手手中。18世纪后期，近代化学之父安托万·拉瓦锡（Antoine Lavoisier）试图在法国制定第一个国民账户体制，他成为包税商集团（Ferme générale）的管理者。这是一个外包海关和收税运营的机构，代表国王征收关税，并加强了旧制度下的税收机制。在法国大革命之后，拿破仑·波拿巴（Napoleon Bonaparte）要求统计机构提供公民的具体信息，以便更有效地征兵，同时加强国家的收税和征用能力，设计更好的战时经济管理方式。第二次世界大战期间，美国商务部编制的国民收入账户被用来系统性地评估罗斯福总统的《胜利计划》，帮助美国在战争冲突中作



出决策。

在当代治理中，国际排名对全球化国家的信誉至关重要。研究表明，诸如世界经济论坛所产生的全球竞争力指数已经加强了新自由主义的做法，这些做法逐渐将政府的角色从公共利益推动者变为了市场扩张的支持者。这一趋势还通过将信用评级纳入国家政策和国际条约得到了进一步强化。各种类型的善政指标，包括以国家有效性、腐败程度和市场开放为重点的指标，对于定义国家间的全球等级制度和国际“黑名单”都十分重要，而从八国集团、二十国集团到经济合作与发展组织（OECD）的建构，以及世界银行和国际货币基金组织（IMF）的权力分配来看，经济绩效指标（主要通过GDP计算）成为筹划全球治理机构的主要参数。

法国小说家奥诺雷·德·巴尔扎克（Honoré de Balzac）曾经说过，社会是按照统计学家的指示来组织的，他们在权力机构中占有重要的地位。通过数据，“社会孤立每一个人，为了更好地主宰他们，它能分裂一切、削弱一切。它统治着所有单位，统治着堆积如麦的数据。”无论数据是用来加强制度还是推动改革，它们影响政治的力度可能都是其他任何社会结构无法比拟的。举个例子，如果我们要测量人口，原则上来说，这并非一个不确定的事件，但事实上，人口统计会涉及重要的政治问题，它会影响选举结果，导致资源分配的变化。在美国，为了改变无家可归之人和双重居民身份的公民代表名额不足的情况，人们已经做了各种尝试。然而，由于每次更正都会影响到特定的司法权、种族或族裔类型以及联邦税收的分配，激烈的争论在历史进程中妨碍了这一问题的改善。詹姆斯·麦迪逊（James Madison）在《联邦党人文集》第五十四篇中写道：

极其重要的是，各州要认识到，要尽量少有增加或减少其人口总数的偏向。如果他们的代表份额是受这条规章管理的，他们就会有兴趣多报人口。如果这条规章只决定他们分摊的税额，相反的诱惑就会占优势。

# 数据与政治

统计与政治（和政治利益）之间的关系历史悠久。统计学的先驱被称为“政治算术”，这是一个17世纪出现在英格兰和法国的学科，主要侧重于测量以城市规划为目的的人口趋势和预期寿命。随着时间推移，这些计算被系统性地整合到公共决策中，从而成为国家政策的重要组成部分之一。“统计学”这个词的词源来自于数据和政府之间的内在关系：它成为利用数据使国家变得强大的一门科学。“国家主义者”即是那些进行数据研究的统计学者。最初的统计学者认为，通过量化社会现象，他们的模型能确定一个基本的社会模式，从而解决社会中一些最关键的问题。在统计哲学的背后，存在一种观念，即社会现实是建立在某种社会秩序的基础之上的，数据能以其强大的准确性揭示这种秩序。现代数理统计学的创立者卡尔·皮尔逊（Karl Pearson）充分体现了“实证主义者对量化十分狂热，并有极大的社会野心”。对皮尔逊来说，世界并不是由实物组成，而由对其的感知构成。自然本身并不具有特定的形式。因此，科学的目标是通过一种明确的方法使自然变得井然有序。在他的巨著《科学的规范》中，皮尔逊主张，“科学的领域是无限的；其可靠的内容是无尽的，每一组自然现象、社会生活的每一个阶段、过去或现在发展的每一个时期都是科学的材料。整个科学的统一仅在于它的方法，不在于它的材料”。

不可否认的是，与其说皮尔逊的哲学能用来理解世界，不如说是用来掌管世界。正如《相信数据》一书的作者，历史学家西奥多·波特（Theodore Porter）所言，量化是“一种社会技术”。算术和代数不像现代数学，根植于古代的几何学，强调理论论证，“与数据领域相隔甚远”。作为统计学的基石，它们“诞生于实用的艺术中”。在现代社会，量化过程在实验实践中起着核心作用。其主要目标之一是“作为实验室的物质文化与形式理论的预言之间的桥梁”。尽管我们大多数人认为这个理论测试角色是数据量化过程的决定性任务，在实践中这却往往是不正确的：“研究人员在处理缺乏数学理论的议题时，会在报告研究方法和定量形式的研究结果上同样一丝不苟，并过滤掉无法明确表达的发现。”

这种类型的推理应用，渗透了维持社会生活秩序的意愿，不可避免地模糊了主体与调查对象之间所谓的距离。在现实世界中，我们衡量事物的方式也难免会对事物本身产生影响。“社会统计学能描述社会；但它同样也是社会布局的产

物”，特拉华大学社会学教授霍埃尔·贝斯特在《该死的谎言与统计数据》一书中写道：

让我们注意到社会统计数据的人有理由这样做；他们难免想得到一些东西，就像重申和公开某些统计数据的记者和媒体人一样，有着他们的目标。统计数据是用于特定目的的工具。若要批判性地思考统计数据，必须了解它们在社会中的地位。

英国广播公司（BBC）节目《或多或少》的前主持人迈克尔·布拉斯兰和安德鲁·迪尔诺将统计数据在社会领域的应用代入了印度寓言《盲人与大象》：

六个印度人

学习很认真

他们去看大象

（虽然他们都是盲人）

每个人通过触摸

满足了想象

我们知道，尽管这群人非常想从感知（而非理论）中获取信息，他们一次也只能触摸到大象的一部分。所以对于只接触到大象身体侧边的人来说，它就像是一堵墙。对于摸到象鼻子的人来说，他认为这是一条蛇。摸到象腿的人觉得它像一棵树，而抓到尾巴的人则觉得它是根绳子。他们会因此“无休止地争吵……尽管每个人都有部分正确/他们却都是错的！”。根据定义，统计数据是静态的。“你要计算的事物必须保持不变，”英国智库新经济基金会（New Economics Foundation）成员和《数据的暴政》一书作者戴维·博伊尔认为，“但现实生活并不是静止的”。

普利策奖获得者，美国人口统计文集《数据政治》的联合主编之一，社会学家保罗·斯塔爾称，数据就像是照片：它们“似乎阻止了人类活动的流向，并因分离的检视而固化”。因为单纯的文字就像是绘画，需要被解释，数据意图呈现一个固有的现实。事实上，“统计数据不仅仅用来做出许多解释，它们其实也包含着这

些解释。因为数据不是简单地再现现实，整个数据系统代表的，是社会生活中的独立因子”。

正如布拉斯兰和迪尔诺指出的那样，问题不只是对于概念的混淆和误读。当统计数据应用于政策制定时，它们会导致各种各样的后果，其中的大部分是我们无法控制的：“当我们背向不测量的部分时，它们会做出异常的变化；就像当我们触摸感知（大象的）腿部时，象鼻子会出现意外的情况”。例如，在医疗保健方面，绩效评估（以及公共资金的分配）通常与一系列有限的指标相关联，如通过手术得救的患者与在治疗中死亡的患者数量之比。在美国以及越来越多的欧洲国家，这给了医生不正当的激励，因为当失败风险太高的时候，医生不愿意接收这些高难病例，而会宁可让患者离开医院（或取消所有预约）。同样，在公立医院对等候时间进行评估来考核的英国，公众调查也发现了管理者的一些计谋、暗中操作和它们造成的不良后果，比如取消后续回访。根据对一家“可疑”医院进行的调查发现，取消后续回访导致了至少25名患者由于缺乏术后护理而失明。

这种基于标准化评估的目标政治已经入侵了公共管理的大部分领域，而主观性和自由性曾是这些领域中最重要的一部分。例如，如今教育从根本上被困在了特定的成就参数中。学校在标准化考试的基础上进行评分，鼓励学生互相竞争以获取公共资金，或是吸引一些有钱的学生（和他们家长的捐款）。2001年美国的《不让一个孩子掉队法案》（NCLB）由一个大的两党联盟推动，可以说真正改善了美国公立学校为贫困儿童提供教育的方式。该法案通过使联邦资金与测试制度和处罚的接受度挂钩，来促进问责制和学校实效，从而改变了美国学校的运作理念。包括教学专业人士、民间社会团体和非政府组织在内的评论家认为，这项法律正在破坏公立学校和公共教育系统为穷人和少数族裔儿童提供教育的能力。在《儿童留守》一书中，一些学者、活动家和教师坚持认为“NCLB是在惩罚而非帮助穷人/少数族裔儿童（和他们的学校）”，希望维持公共教育私有化的议程，因为“对测试的关注和测前准备降低了校内课业的难度”。教师网在2007年进行的一项调查显示，对于大多数教师而言，在重要测验中对NCLB的强调并不奏效。只有37%的受访者认为标准化测试“有点用”，而42%的受访者则认为这对他们的教学“完全不起作用”。此外，超过40%的学生声称这些测试是在鼓励他们死记硬背，44%的学生认为标准化测试会让他们忽略那些考试不涉及的课程材料。美国公平与公开考试中心表示，标准化测试是在“鼓励浅显问题的快速应对”，“并不衡量在任何领域进行深入或创造性思考的能力”，“试图缩小课程范围，采用过时的

教学方法和有害的练习方式，诸如记录和跟踪学生的成绩”。在2012年底，奥巴马政府承诺修改和完善这项政策，并给予了三十多个州豁免权。

博伊尔指出，“因为衡量真正重要的东西非常困难，所以政府和机构会试图明确定义一些其他的東西。他们必须这样做。但定义一些错误的东西会造成非常严重的后果。”他提到了英国的学校排名表，排名中将标准化考试的超然力量渗入了学校间竞争：“麻烦在于，每个学校会集中注意于测试结果，以提高他们在表中的排名。这也意味着学校可能会放弃给成绩拖后腿的学生，把注意力集中在那些成绩是D级的学生身上，他们是唯一会在考试成绩榜上起到重要影响的人，而其他人则会得不到应有的关注。”他总结说，“如果你选择了错误的测量方式，有时就会得到与所想相反的结果。”

## 相信数据

要在公共生活中建立有效的测量手段，需要采取一定程度的强制措施。例如，要在建筑领域执行共同标准，既要求私营企业遵守规定，也需要一支训练有素的劳动力队伍。同时，数据也是强有力的说服武器。借用法国哲学家米歇尔·福柯（Michel Foucault）对政府行使权力间接手段的分析，人们可能会认为，评估的“客观”标准构建不仅使“其他人的监管或控制成为可能”，对“自我身份”的塑造还会“通过个人的自我认知和自我调节进行”。简而言之，由数据制定规则比单纯的自上而下强制要复杂得多：它包含了很大程度上的自满情绪，并从根本上设计了主体的行为方式。总之，这是一个自愿默从的体制。因此，它成为福柯确认政府合理性或“政府性”的一个关键方面——也就是一项“引导行为”和从“一定距离”影响行为的技术；或者说，是按照“具体的结果”和“多种策略”来“处理事物的正确方式”。

数据也是界定现代官僚治理体系运作方式的基础。新的知识模式确实对决策者的体制机构的定义至关重要，决策者的选择必须由除行政自由裁量权之外的参数来指导。现代官僚制研究的前驱马克斯·韦伯（Max Weber）认为，官僚主义的本质是技术的力量，它导致了所有与政治斗争相关的非理性和情感因素的边缘化，也就是人为因素逃脱了精确的计算。韦伯的想法强调“公共行政秩序能够实现距离、合理性、客观性、权威性，以及计算机制的方式”。历史学家西奥多·波特一直在研究我们社会对数据信任的历史演变，正如他所言，量化的演进过程与官僚政治的发展有着内在的联系。出于法律和政治的原因，行政自由裁量权被人怀疑，“所以监管者别无选择，只能不懈寻找事实，并尽可能地将它们减少到几个决定性的数据”。特别是在像美国这样的多元民主国家，利益集团争夺在公众中的地位以及参与决策的时候，基于统计的决策有助于调整政府、企业和整个社会的利益：“当价值观的冲突与共识难以捉摸，数据和操纵它们的技术会因为它们表面上的中立性而被推崇。有了统计数据，激烈争论的问题看起来似乎都会被马上解决”。

在20世纪30年代的大萧条时期，美国统计学家和经济学家发现他们在公共决策中起到的作用呈现了指数级的增长。一系列的环境灾难，特别是1927年的密西西比河大洪灾，为基础设施项目大量使用基于统计的风险评估方法创造了良好的

条件，同时也为评估预防政策的成本和效益提供了初步研究。随着经济衰退，公共预算不断缩减，“有限”的资源配置中引入了新手段的指导，以避免政治争议。1936年，《防洪法》首次在公共政策中引入了成本效益分析，指出除非有防洪计划能证明其效益超过成本，否则不会有任何计划获得联邦资助。最初，成本效益方法的设计是为了鼓励高度争议部门的开放和中立，例如防止环境灾难及其与公共基础设施项目和工业化的关联。当然，数据的可信度会在专业性、科学中立性和透明性的修辞中体现出来。在1936年《防洪法》中设置成本效益条款的参议员认为，负责这些项目计量经济学分析的专家是“值得尊敬的、直率的、爱国的人”，因为新的评估体系需要“按照轻重缓急，做出独立的、非政治性的、没有偏见的决定”。由于这些人的专业知识和声誉，负责处理数据的部门开始享有越来越大的影响力，这种影响力又反过来被用以增进与强大的私有利益及企业的友好关系。罗斯福政府内政部长哈罗德·伊克斯（Harold Ickes）表示，负责量化管理发展的官员是“华盛顿最有力、最雄心勃勃的游说团体”，与军事分裂和“官僚制下最完美的部分”紧密联系在一起。

在20世纪60年代和70年代，用于衡量成本和收益以及评估公共基础设施项目有效性的各种模型，已经“从地方官僚体系中聚集起来，变成一套合理化的经济原则”。在美国人对政府“经常性不信任”的政治背景下，统计分析的拥护者在各个领域都引入了这些工具，包括从社会福利项目的评估，到监狱的运作，等等。他们声称，这些工具在公共政策的计量经济评估上“几乎普遍有效”。特别是，成本效益分析和实验测试程序开始由各种下属机构和私营公司实施，这些机构和公司会将他们的专业知识出售给政府部门或招标公共项目的公司。正如波特强调的那样，在这个阶段，“科学、国家与大型行业的利益之间”存在明确的一致性。例如，尽管前美国总统尤利西斯·格兰特（Ulysses Grant）于1877年向国会作出了一个著名的不成功呼吁，但美国国家标准局经常鼓励政府和有关工业部门进行密切合作。格兰特在呼吁中对私营部门的检测和评估与国家要求的不同：“这些试验不能由私人公司自身执行，不仅有费用的原因，还因为结果必须依赖无私利的人的权威。”

在20世纪90年代，量化评估的新做法成为“审计爆炸”的一部分。除了通过财务审计对私营公司会计进行监管外，还出现了“环境审计、效益审计、管理审计、法务审计、数据审计、知识产权审计、医疗审计、教学审计和技术审计等评估方式，它们带来了不同程度上的制度稳定性和接受度”。《审计学会》一书的作者迈

克尔·鲍尔（Michael Power）认为，会计制度进一步强化了人们对数据的关注，并铸就了“理性化社会的神话结构”：“审计爆炸的根源在于组织生活的程序性重组，以及新的‘治理合理性’。”

这样的重组给所谓的专家，也就是生成数据的个人和组织，产生了前所未有的影响。用德国心理学家格尔德·吉仁泽（Gert Gigerenzer）和他的同事们——《机遇帝国：概率如何改变了科学和日常生活》一书作者的话来说，专家的权威性因为统计和计量经济计算结合到公共决策中，而得到了极大的提升：“事实上，数据的爆炸创造了一种新的专家，他们的主张更多依靠信息与形式技术，而不是具体的经验或个人判断。毕达哥拉斯赋予了数据如此大的影响力，这对于新型专家来说既是一种福音，也是一种诱惑。”把这些“新”专家的工作动员起来的标准化程序“与现代官僚制的公正性和规则相平行”，它们试图排除“个人自由裁量权，强调既定程序的一致性和机械性应用，以避免偏见，一个针对事实，另一个针对公平”。

然而，大部分的审计报告并不传达有价值的信息。对于外部用户来说，它们不是基于证据的、不需加以说明的文件。总的来说，这些报告是自成一体的不透明记录，从根本上依赖着“中立，客观，冷静，专业”的语言系统。这实际上意味着对公众的全面披露、透明度和问责制是通过专家认证被限制的。审计过程因而成为一条捷径，它建立在我们社会对专家根深蒂固的信任基础上，而非理性的公共审议基础：“这是问责制链条中的一个死胡同。”简言之，更多的数据和会计制度并不一定等同于更好的问责制。相反地，它们会减少“公众猎奇的心理和公开调查”，因为一个必要事实是，专业审计的最终用户并不是广大公众，而是“专家话语中的虚拟参照点”。这也是计数的很多悖论之一。虽然审计爆炸发展是以问责的名义发生的，“给出账目看起来却像是一种避免问责的方式”，在这样的情况下，数据扼杀了政治讨论和社会讨论，意图提供不容置疑的事实。自相矛盾的是，“虽然审计学会宣称的纲领性基础是开放和问责，它却有着日益封闭的危险。”

使用数据的专家成为社会信任的守护者。数据和守护者的权力从根本上颠覆了支撑社会和政治联系的委托代理关系。公民、选举代表和其他利益相关者（委托人）被专家（代理人）所控制。而基于数据的治理机制的深刻制度化进一步加强了数据的力量，因为“我们促成了各种辅助认证或可靠性担保……这些很容易被操纵，但现在对于那些已经不怀疑新代理人的委托人来说，又是非常重要的东



西”。我们都相信那些以数据为证据的人，即使我们知道，各种目的的数据造假都是如此容易。在学术研究领域，每年都会发现数百起数据篡改案例，甚至涵盖了世界一流大学。当然，它们之中的大部分并不会成为世界性的头条新闻，但确实有一些引起了轰动。比如说，在1986年，诺贝尔奖获得者、生物学家戴维·巴尔的摩（David Baltimore）就因为卷入了一项涉嫌捏造数据的研究，引起了全球热议。这些造假数据被用以支撑一项免疫系统研究中的新发现。尽管巴尔的摩最后摆脱了所有控告，他还是在这项惩戒中受到了一定打击。在社会研究中，也有非常多相似的例子。历史学家迈克尔·贝勒斯雷斯（Michael Bellesiles）的例子可谓臭名昭著，他伪造了一份“独特的”数据集，其中包含了数百年来美国民众中小型武器的分布情况。他写作的畅销书《武装美国》被《经济学人》杂志称颂，并获得了久负盛名的班克罗夫特奖。而后，学术界发现了书中的数据造假，这也迫使贝勒斯雷斯放弃了奖项。

哈佛大学经济学家卡门·莱因哈特（Carmen Reinhart）和肯尼思·罗格夫（Kenneth Rogoff）是《这次不一样：800年金融荒唐史》这部金融危机史上影响深远的著作的作者。2010年，他们发表了对美国国家经济研究局（NBER）的后续研究报告。该报告很快成为所有鼓吹欧洲和美国紧缩计划的政府和保守势力的参考。欧盟委员奥利·雷恩（Olli Rehn）和美国共和党议员保罗·瑞安（Paul Ryan）都引用了这项研究。莱因哈特和罗格夫题为《债务时代的增长》的论文使用了“44个国家长达200年”的时间序列数据和“3700份年度观察报告，其中涵盖了广泛的政治体系、制度、汇率安排及历史情况”。其主要实证结论为，各国的政府债务不得超过GDP的90%，否则会自动引发经济增长速度的变缓和系统性衰退的风险。之后，在2013年，马萨诸塞大学阿默斯特分校的一名年轻学生被指派完成一项作业——重复一篇经济研究领域著名论文的运算过程。他选择了莱因哈特和罗格夫的这篇论文，因为每个人都熟知它在经济危机的政治管理上具有的影响力。然而，在经过多次尝试后，他发现无法重复论文中的运算。他的导师建议他联系原作者，他最终收到了原始电子表格，发现了许多基本的计算误差，包括错误的平均值，这在很大程度上破坏了研究结果。莱因哈特和罗格夫为他们的错误公开道歉，但他们的总体结论仍然得到了部分支持，而凯恩斯主义的经济学家则抨击他们是错误的决策者。美国经济和政策研究中心批评了这项研究的结果，以及它为在美国和欧洲引起广泛不满的紧缩政策所提供的辩护。该研究中心反问道：“莱因哈特和罗格夫的算术错误造成了多少失业？”伦敦大学经济学教授丹尼尔·哈默梅什（Daniel Hamermesh）怀疑，由于这项研究在政策上的应用，可能直

接导致了失业现象：“但它为那些影响了人们如何看待这个世界的事物提供了一种知识上的合理化。而人们，特别是政治家，他们看待这个世界的方式会最终影响到世界的运行”。

两位知名学者怎么会犯这样的低级错误，还在如此长的时间里侥幸没被发现呢？他们怎么发表在美国国家经济研究局（NBER）的工作论文中呢？要知道，NBER可是世界上最负盛名的经济智库，因包括了22名诺贝尔奖获得者而引以为傲，但却没有人注意到论文中错误的平均值计算。美国国内外的报纸，以及像国际货币基金组织这样，在两位经济学家加入学术界之前曾经工作过的机构，都褒扬了两人的成果。为什么没有人发现这样严重的计算错误呢？当然，这个事件也严肃地质疑了所谓的同行评审过程的可信度，我们不禁怀疑，还有多少影响着日常政策的研究中，也存在着类似的“偏爱”。（本书分享更多索搜@雅书）

这样的问题不仅仅在美国发生。在欧洲，社会心理学家迪德里克·斯塔佩尔（Diederik Stapel）是很多被广泛引用的论文的作者，他的论文经常发表于像《科学》这样的著名刊物上。他在2012年承认自己曾“修改了研究数据，伪造了某些科研项目的结果”，“不止是一次，而是好几次”，“不仅是在一小段时间内，而是持续了很多年”。可以说，他的整个研究生涯，那些曾在社会定性研究领域给教育计划和政策带来很大影响的研究，都是以杜撰的数据为基础的。在斯塔佩尔的一本回忆录《出轨》中，他用强大的叙事力描述了他篡改数据的过程：

我更愿意在家中做这些，在深夜，所有人都睡去的时候。我会给自己泡一壶茶，将电脑放在桌上，从包里拿出便笺，用钢笔写下我将要完成的一整套研究项目和成果……后来，我开始输入我臆想的数据，一行一行、一列一列……3、4、6、7、8、4、5、3、5、6、7、8、5、4、3、3、2。输完数据后，我就可以开始做第一项分析。这些数据通常不会立即产生正确的结果，于是我便返回模型并修改数据。4、6、7、5、4、7、8、2、4、4、6、5、6、7、8、5、4。直到所有分析结果都能按照预期实现为止。

斯塔佩尔坦诚地向所在大学的调查委员会承认了这一切：“我没有经受住想获得成功、想发表论文、想成为更厉害的学者的诱惑……在一个很少有制约平衡、人们总是单独工作的制度下，我走错了路。”

## 数据、市场和民主

苏格兰社会评论员托马斯·卡莱尔（Thomas Carlyle）在《宪章运动》一书中写道，一位机智的政治家“可能会用数据来证明所有事情”。我们都知道一句著名的格言：“世界上有三种谎言，即谎言、糟糕透顶的谎言和统计数据”。用吉仁泽和他同事的话来说，如果我们的社会已经意识到数据能够用来证明什么，“那是因为他们经常做了（错误的）证明”。当专家的主张过多时，“政治和社会制度允许他们这么做，甚至鼓励了他们的虚荣自负”。法国哲学家雅克·埃吕尔（Jacques Ellul）曾经写道，“现代人需要一种与事实真相的联系”，这像是一种“用某种方式说服自己的自我辩白，让他觉得自己是一个服从理性和实证经验的人”。但埃吕尔的所指并不是科学，他的分析重点在于宣传的意义和形式。在埃吕尔的理解中，数据的使用对于“在理性和事实性因素的基础上创造非理性的反应”是至关重要的。

为了成为社会辩论的一部分，个人和团体必须呈现一些数据。没有数据，论据将被视为缺乏可信度，而只是基于纯粹的经验之谈。正如贝斯特所说，“数据是因为要为政治斗争提供武器而被创造和复制的，而这种政治目的通常被隐藏在一种断言后面，即数据，仅仅因为是数据，所以一定是正确的。”媒体进一步放大了对数据的操纵。它们迫切需要简单的故事情节。它们努力地汇报“事实”，而粗浅的数据是强大的营销工具，因为它们吸引人们的眼球，容易引发争议，还简化了记者的工作，记者不再需要仔细分析各种意见和细节。正如哥伦比亚大学新闻研究院的一位讲师认可的那样，“我被培养成了这样一种人，我只相信能够观察到的和可以量化的事实……新闻记者在看到平均成功率、股市行情、人口普查、选票统计时，才能得到最大的安全感。”

《数据是靠不住的》一书作者查尔斯·塞费（Charles Seife）在书中引用了美国国防部在越南战争期间发明评估体系的例子，该体系用先进的计算机向媒体提供各种统计资料，以表明美国正逐渐在这个东南亚国家的战争中赢得胜利。但是，事实当然不是这样的：“这种信息是新闻的原始材料。它给了新闻记者们表达观点的依据；如果记者的言论不能被硬性事实支撑，他们在表达时就容易显得软弱无力。”一旦某个数据出现在了新闻报道中，它就走上了自己的轨迹，并会经历一段像洗钱一般“数据合理化”的过程。其可疑的起源会被瞬间遗忘，通过不断地

复制，该数据开始被视为一个明确的事实，“准确而权威”。不久，数据所经之路变得模糊不清：“人们忽略了数据的原始来源，却认为这个数据一定是正确的，因为它随处可见。”

正如美国国防部伪造数据支持保守派媒体的宣传，使媒体得以引用战争中的死亡人数、俘获的武器数量和部队人数来“描绘战争将胜利的美好画卷”。反环保主义者的游说团体也进行了相关研究并呈现出数据模型来佐证各种各样的虚假联想，他们说高油耗的多功能车（SUV）比混合动力车更利于环保，还说全球变暖现象并不存在。数据推理的应用经常会为一种系统性的否定做好准备，强大的利益集团十分支持这种应用，尤其是在美国。数据“谎言”的产生看起来是违反直觉的，它并不是为了说服意识形态的敌人，而是为盟友提供一个数据库。2007年，一个市场研究小组发表的研究显示，众所周知的高油耗车型悍马H3比丰田普锐斯更加节能（他们认为悍马比普锐斯拥有更长寿命和可行驶更长里程，利用这两个因素设计新的模型，降低了前者对整体环境的影响，从而得出结论），反环保主义专家和全球变暖的怀疑者由此重新得到了激励。这些虚假证据迅速席卷了主流媒体，《华盛顿邮报》为气候变化否认者开放了专栏文章，他们呼吁人们购买悍马车，“以代替丰田普锐斯”。正如历史学家内奥米·奥利斯克斯（Naomi Oreskes）和埃里克·M.康韦（Erik M. Conway）在医学和环境研究领域对反对科学的力量进行分析时所表明的那样，媒体已经成为对虚假数据心满意足的支持者，毕竟新闻记者们总是将统计数据视为一种无可争议的证据。考虑到记者是经过培训的持中立观点的代表，数据的呈现已经成为那些有意传播怀疑观点，阻碍某些关键领域改革的人手中的有力武器。比方说，他们会进行风险炒作，将预防原则应用于健康领域和环境规制：“无论错误一方的观点多么荒谬可笑，或者多么依赖‘数据的骗术’，新闻媒体都会报道它、增强它，让这些人造的‘事实’显得很有生命力。”

新闻报道对股市指数的迷恋是媒体与数据间密切关系的最好例子。在20世纪80年代以前，只有专门的媒体机构会提供国内外市场的股票信息，几乎没有媒体会参考股票指数。例如，称得上是世界上最知名的股票指数的道琼斯工业平均指数（通常被称为道琼斯指数），是在1896年首次公开发布的；尽管20世纪的投资者们十分青睐道琼斯指数，它在大萧条时期还曾一度暴跌，它真正流行起来还是在20世纪90年代，那时的综合性报纸和电视台才开始在定期出版物和播放中引入股市信息。同样地，在20世纪50年代开始正式编制的标准普尔500指数，在21世纪才有了全球知名度。纳斯达克综合指数在1971年开始运作，到了20世纪90年代

末，随着互联网泡沫的出现，它才登上了全球的头条新闻。从那以后，这些数据（连同国家和地区的各项指标）几乎成为每个国家公众辩论中的主角。媒体每天都会根据股票市场活动的平均值和估计值，为我们提供股票行情表和时间序列。各种专家会对此作出评论，数据不断在后台滚动。动画图形和复杂的表格给这些数据增加了重要的光环。媒体经常会带给我们这样一种印象：股票指数是一项公益事业，是一国经济健康的指标。交易量比较低的时候，国家就经历着糟糕的一天；交易量高的时候，媒体则会为此而庆祝。这种看法也影响着我们的整个社会：当指数上涨时，我们欢呼；当它下降时，我们忧伤。但股票指数代表的究竟是什么？事实上，这些数据和公益事业毫不相关，仅仅只是用来描述私人市场交易的一部分主体的情况而已。股票指数只包括了那些上市公开交易的公司，它们实际上只是全球私营部门中的一小部分。而且，这些指数只涵盖了这些公司中规模最大的一部分（以总体股票价值计算，而非资本或劳动力规模）。打个比方，道琼斯指数仅仅包含了三十只股票，虽然它们已经能粗略地代表四分之一的总股票市值，但也确实没有包括任何中小型企业。它仍然是美国最主要的市场晴雨表。标准普尔500指数对市场的整体覆盖面比较广，但由于它有市场资本加权制，所以会给大公司一定的特权。最后兴起的纳斯达克指数则只关注科技股。所有的股票市场统计数据也是如此。由于这些指数能被用来吸引投资，它们会攫走本将投向小公司和地方经济的资金。在世界很多地方，实体经济非常有可能受到对股票指数盲目乐观的影响。这些数据并不是经济发展的信号，而是以牺牲中小企业为代价加强大资本，对市场造成的扭曲。

这些指标都不是市场动态的真实参考，更不用说能用以衡量经济健康。然而，媒体炒作已经成功将它们融入了我们的社会心理，从而加强了金融资本主义的政治控制。韦伯将数据的力量与资本主义意识形态的霸权主张联系在了一起。他把资本主义制度定义为“一种附带资本会计的体制，即按照现代簿记和收支平衡原理计算确定收入的体制”。而且，数据的受欢迎程度也影响了我们对价值的理解。股市指数并不能衡量企业的实际价值。它们反映的是股票预计的交易价值，也就是投资者愿意为购买股份花多少钱。当投资者对其非常有信心时，这种价值会大幅上涨；若投资者破产，股价也会相应地崩盘。《哈佛商业评论》前任编辑乔尔·库尔茨曼（Joel Kurtzman）曾强调，“机构如今不再购买和持有很多股票，因为他们信任公司的潜在价值”。恰恰相反，他们会买进卖出股票，通过“一些数学公式的帮助”，缩短持有股票的时间。正如国际政治经济学家罗内·帕兰（Ronen Palan）所说，“我们用来反映股票估值、利润和财富的数据不再能折合成‘真正

的’有形商品，它们是将‘权力’当作纯粹的关系来衡量。”然而，由于媒体自身的数据自信和批判性分析的缺失，这些指标传达了一个错误的信息，即当数值上升的时候，实际财富正在被创造。从而，当市场整体情绪高涨时，我们会感到富有；当这种“虚假的”财富蒸发时，我们突然感到了贫穷。正是因为这些数据，我们都成为市场社会中无意识的利益相关者。

有些数据在本质上显得比较微妙，所以会特别隐蔽。它们不以传统的统计格式显现，也不出现在复杂的公式中。它们非常简单明确，我们几乎忘了它们也是数据，是抽象的概念和人类的发明。价格就是一个完美的例子。我们的日常被价格包围，我们会通过商品标价来看世界，让它帮助我们做出决定。无论是选择假日旅游的目的地、孩子要就读的学校，还是考虑看望父母时，价格都会成为我们做选择的基本参数。价格已经成为价值最有力的标示。每件物品都要值得它们的标价，或者与我们愿意付出的钱对等。正如奥斯卡·王尔德（Oscar Wilde）在《道林·格雷的画像》中写到的那样，“如今，人们知道所有东西的价格，就是不知道它们的价值”。

价格不仅在我们的日常生活中普遍存在，还具有宏观经济统计的功能。例如，有史以来最强大的指标，万能的国内生产总值（GDP）就是以市场价格度量的商品和服务的总和。我在前一本书《GDP究竟是个什么玩意儿》里谈到了GDP背后的政治性。我们在此不多谈，只将GDP视为一项“总额”，而不包括生产过程中的物资贬值（比如机械、工具、车辆等）。在市场之外的各种交易（例如家庭内部交易、非正式经济、易货交易等）也不被计入总额。此外，GDP并不会考虑到经济增长过程中消耗的自然资源价值，因为这些资源能免费从自然界获取。GDP甚至也不会将污染和环境退化的经济成本纳入，而这是工业发展造成的明显后果。所有这些重要的遗漏使GDP对经济实绩的衡量具有很大的选择性（有些人可能会认为这一数据会使人变得目光短浅）。比方说，家庭服务即使没有在市场上被定价和交易，也仍然会产生根本性的经济影响力。如果政府不得不为家庭层面的各种服务付费（包括对儿童和弱者的照顾、教育问题等），那我们的经济发展就很可能陷入停滞状态。最近，一项估算美国家庭生产价值的研究表明，从1965年到2010年，美国家庭内进行的各种生产活动每年占经济产出的30%以上，这一比例在1965年达到了39%的最高值，2010年时则下降到了25.7%。在很多国家，“临时工作”和非正规的商品和服务交易为数百万人提供了必要的生活来源。尽管这并不会在GDP上表现出来，却常常构成了实体经济的骨干。同样地，因为

自然资源没有定价就忽视了它对经济的投入，这使我们遗忘了只有在生态系统不断提供“资本”的情况下，经济增长才能得以实现。没有干净的土壤、水、空气和其他重要的天然过程，农业生产是无法进行的。没有地球提供的化石燃料、碳氢化合物和能源，工业化也无法实现。然而，当这些资源枯竭时，我们面临的风险不仅是经济发展的停滞，还有危及生命的自然失衡。会计学基础告诉我们，利润等于收入减去“全部”成本。由于GDP经常忽略经济中的一些关键部分和重要的成本花费，没有任何明智的商人会将它应用于公司经营。然而，它已经成为整个社会运行的重要参数，正如一篇发表在《经合组织观察家》上的文章提到的那样：

如果说统计界有一个颇具争议性的标志，那一定就是GDP。GDP衡量的是收入，而非持平状态；它衡量增长，而非破坏。它会忽视社会凝聚力和环境的价值。然而，政府、企业，还有大部分人，都无比坚贞地信赖着它。

定价机制在制定公共政策方面的应用有着悠久的历史，它可以追溯到欧洲和美国现代工程技术发展的时期。当时，在基础设施的建设中，人们需要数据化的评估工具，以便给税收征管带来便利。19世纪下半叶，含有量化因素的公共管理变得相当普遍，尤其是在公共工程的定价问题上、招标的分配，以及公民在通过大桥和铁路时需支付的通行费的计算，都会用到定量的方法。举例来说，法国工程师朱尔·杜普伊（Jules Dupuit）就是第一个在火车票定价中引入诸如边际效用递减率等概念的人。杜普伊认为，数学推理的确定性对于良好的政治经济分析至关重要。他解释道，立法者的准确定位应该是要“把这些政治经济所表现出来的事实奉为神圣”。

弗里德里希·奥古斯特·冯·哈耶克（Friedrich August von Hayek）是自由市场经济学的先驱，他率先将价格综合理论发展成为各种指标，也就是信息带来的信号。他曾于1945年在《美国经济评论》期刊上发表了一篇颇具影响力的文章，名为《知识在社会中的利用》。他在文中写道，价格体系是“一种记录变化的工具，或一种通信系统，这种通信系统能使单个的生产者像工程师观察一些仪表的指针那样，来观察一些指标的动向”。哈耶克认为，社会是“一个关于相关事实的知识分散地掌握在许多人手中的体系”，而在这个由部分信息构成的万花筒中，“价格能协调不同个人的单独行为，就像主观价值观念帮助个人协调其计划的各部分那样”。然而，尽管哈耶克对价格的交际功能抱有极大兴趣，但他也认识到价格只是“一种数据指标，这种指标不可能从特定的物品所拥有的任何特征中得出，但它

却可以反映，或者在它身上集中了它在整个方法目的结构中的重要性”。

在哈耶克将价格定义为信息的信号后，经济学家费雪·布莱克（Fischer Black）、迈伦·斯科尔斯（Myron Scholes）和罗伯特·莫顿（Robert Merton）开发了复杂的公式来预测各种金融交易中的价格，特别是在衍生品市场。三人建立了布莱克-斯科尔斯-莫顿模型（Black-Scholes-Merton model），莫顿和斯科尔斯因此获得了1997年的诺贝尔奖，这一模型很快就成为全球金融领域定价方法的重要参考。随着世界经济日趋金融化，价格成为价值的完美替代品，金融市场成为资源配置的可选择空间，这也进一步影响了我们的治理模式和社会评估财富的方式。

直到2008年9月，这个有着完美定价规则的世界开始崩溃了。美联储前主席艾伦·格林斯潘（Alan Greenspan）在众议院监督与政府改革委员会的一场听证会上承认，“定价模型的发现曾被授予诺贝尔奖，这个模型奠定了金融衍生品市场的发展。现代风险管理思想盛行了数十年，但是这个知识体系却已分崩离析。”那只是全球金融危机的开始，而后，这场危机席卷了整个世界，并造成了大萧条以后最具破坏性的经济衰退。然而，价格对我们社会的控制是不受金融世界的经济混乱影响的，在经济危机阶段，数据的力量比以往任何时候都更加强大。

价格是指标。它是在资源稀缺的情况下，因供求关系而确定的综合性参数。从理论上来说，价格应该大致对应于商品和服务给个体消费者带来的边际效用。实际上，在大多数实体经济体中，价格根本不是（或只是部分）由消费者的偏好和优先选择顺序确定的。一般来说，价格很容易受到供求关系良性的外部动力因素的影响，其中包括税收、补贴、企业联合、垄断等，更不用说能够提供给某些行业特惠待遇，却不适用其他行业的鲜明政治立场。由于价格（和货币量）受信贷影响很大，金融家、投资者、银行家和保险公司的贴现率在很大程度上决定了什么样的东西才是真正有价值的。我们的经济状况比以往更依赖少数人的偏好。

价格和市场总是齐头并进的，因为前者是交换价值的具体指标。显然，那些附有价格标签的物品就能在市场上进行买卖。因此，价格的普适性导致了市场的增长远远超过了惯例上与利润相关的传统市场界限。我们现在制定了碳的价格，并有着与碳排放相关的市场；我们有适用于非盈利部门的财务模式；用市场和价格来迎合生物多样性，应对碳排放和森林资源管理；根据保单的价格来评估一个人的生命价值。诸如成本效益、支付意愿、重置成本和投资回报等经济概念在我



们的治理模式中占据了主导地位。从局部上来说，大多数的公共机构都采用了从商业中借鉴的管理公式；在全球范围内，我们在气候变化和环境退化等问题上引入了市场机制。最近的一些研究也显示了数据化组织的运作原理，即通过数据的合成表现出某种财富，以促进资本在离岸经济中的流动，这样就会加大超级富豪和其他人之间的全球性差距，也破坏了各国维护福利制度的能力。

2004年，哈佛大学校长劳伦斯·萨默斯（Larry Summers）公开庆祝了数据的胜利。他称赞“魔球”已经成为棒球界最成功的预测模式，并坚持认为，“在棒球运动中正确的部分，实际上在广泛的人类活动中也同样是正确的”：

在过去的三十年里，投资银行领域已经由善于在第19洞（高尔夫术语，指高尔夫会所中的酒馆）接待客户的人主导，变为由善于解决衍生品定价中复杂数学问题的人来主导。环境监管领域在实际应用中已经基本不再受执行活动的人士或是律师的指导，而让位于能够熟练地进行成本效益分析的人。在总统竞选活动中，组织者曾经呼吁一批有才华的律师来进行辅佐，但现在他们更需要出色的经济学家和工商管理硕士（MBA）的帮助。我可以提出很多这样的例子。这正表明了从社会科学发展而来的分析技术正在得到越来越广泛的应用。

全球市场的“转变”由于对数据不受约束的信任而得到加强。正如亚当·斯密对“看不见的手”的描述，也就是对市场效率在分配资源和解决复杂的分配问题上所做的创立性假设一样，现在大家会普遍“相信数据最终能够‘自动’起作用，可以这么说”。但其实数据自身能做的非常少。它们能藏起冲突，却不能解决冲突。它们只能通过对无法逃避的问题的外在表现分析来遮掩问题本质。毫无疑问，数据是知识进步和治理改善的重要工具。而同时，它们也会成为那些希望维持现状的人手中的有力工具。

---

## 第二章

# 新的全球统治者：信用评级成为不可控制的力量

---

“在我看来，如今的世界有两个超级强权，一个是美国，一个是穆迪。美国可以用炸弹摧毁一个国家，穆迪可以凭借信用降级毁灭一个国家。有时候，两者的力量，说不上谁更大。”

——托马斯·弗里德曼

“当标准普尔或穆迪发声时，那并不是‘市场’的声音。这种声音来自一些掌控评估过程和履历糟糕的人。我们不知道他们的活动会有多大的影响。”

——保罗·克鲁格曼

**信**用评级是全球事务中最强大的数据指标之一。它们涵盖的范围几乎没有限制。企业、银行、保险公司，甚至主权国家都需要评级才能运转。虽然它们通常用字母与数据混合的字符表示，但评级与其他像指标或指数这样的测量手段并没有区别。实际上，在很多方面，评级是赋予数据全球性力量中最独特的例子。此外，自从全球金融危机爆发以来，评级开始成为决定国家政治、经济的数据万花筒中的焦点。利益冲突、技术统治论者的权力和市场的影响在评级中如此明显。事实上，并没有其他的领域会出现这种情形。我们一起来思考一下，在2001年，能源巨头安然公司（Enron）破产的五天前，主要的信用评级机构仍认为它的债券信用为“投资级”。在2008年，投资银行雷曼兄弟（Lehman Brothers）申请破产保护的前三天，其债券仍然获得了主要评级机构的最高级别。甚至在其破产的当天早上，信贷分析师也认为雷曼兄弟值得投资。保险巨头美国国际集团（American International Group）和华盛顿互助银行（Washington Mutual）在倒闭之前一直保持着优质评级，倒闭时发动了大规模的公共紧急救助。美国国会的一份报告发现，评级机构没有动机“给那些一度增加收入，提高股价，扩增高管薪酬的证券指定更严格的信用评级”。

2011年1月，美国政府任命的金融危机调查委员会报告称，主要的信用评级

机构是“破坏金融车轮的关键齿轮”：

处于危机中心的抵押贷款相关证券在没有批准的情况下不可能上市销售。投资者往往盲目地依赖信用评级机构。在某些情况下，他们被迫运用评级，否则监管资本的标准就会随之改变。没有评级机构，这场危机就不会发生。评级使市场猛涨，而2007年和2008年的降级对市场和企业造成了严重的破坏。

然而有趣的是，评级机构在之前就已经对其评估的不准确性做出了警告：“任何采用机构信息的用户都不应该在做投资决定时依赖信用评级或此处的任何其他评价”，世界上最大的评级机构会在它发布的所有报告底端做出这样的免责声明。

但信用评级机构究竟是怎样的存在呢？它其实是“评估企业或政府发行的债务证券（债券和其他证券）的组织，并会根据偿还债务能力对这些证券赋以‘信用评级’。”这些组织会评估债券发行人的信誉等级和投资风险，即违约或者偿付违规的可能性。它们会定期发布报告，将债券分为从AAA（顶级信用）到D（违约）的一系列等级，这为用户提供了可比风险评估，克服了金融市场的信息不足问题。由于公司（即借款人）总是比外部投资者（即放款人）拥有更多的自身财务状况信息，评级机构就会作为中介，试图通过考察借款机构的具体细节来弥补这种“不对称性”，做出呈现公司财务可信度的评估报告。风险水平决定了投资的利率，从而决定负债成本和债务人获得新投资的机会。此外，由于某些国家对限制投机级债券投资有一定的规定，评级也决定了某些机构投资者的证券组合的合格性。同样，监管机构会运用信用评级来确定保险公司所持储备是否充足。

因此，这些机构会对全球资本的流动产生巨大影响，并不可避免地将这种影响投射到全球治理中。目前，三大评级机构（即所谓的“三巨头”），美国标准普尔公司（Standard & Poor's，或S&P）、穆迪投资者服务公司（Moody's Investor Service）和惠誉国际信用评级有限公司（Fitch Ratings）控制着全球95%以上的评级市场。在华尔街崩溃的那一刻，标准普尔公司控制着大约40%的市场，穆迪控制着约39%，惠誉持有另外的约16%。由于惠誉主要在专业市场运作，穆迪和标准普尔在全球评级部门有效地形成了双头垄断。这种市场准入的倾斜分布使任何有意吸引外国投资的大公司或政府都需要通过这些机构的批准才得以运作。评级机构是政治学家科林·克劳奇（Colin Crouch）所谓的“巨头公司”一个很好的例子：它们都是跨国组织，在世界各地都有分支机构，而且它们给出的数据占据了市场的主导地位。由于在全球层面上的监管缺失，这些公司能够制定自己的标

准，从而决定全球经济的管理规则。

评级机构是全球治理中的“怪兽”：它们是具有公共目的的私人公司，“因此它们被称作信用评级机构（agency），而非信用评级公司（firm）”，但在所有权、雇员和收益方面，它们完全是私人化的。一些撰稿人会将评级机构的监管角色看作是正在进行的“世界政治私有化”的明显例子。因此，这些公司的私人治理功能可以被解读为权力中心转变的长期趋势的一个标志，特别是在全球经济领域：私人行为者成为“真正的参与者”，而国家的权威持续下降，自主权减弱。在一些人看来，评级机构应该被“更恰当地视为准政府实体”。而对另一些人来说，评级过程产生了一种“无政府治理”的制度，使这些强有力的数据制造者成为“实际上的全球公共政策私人制定者”。无论人们赞同怎样的具体观点，它们的力量显然会削弱国家和其他公共机构实施公共政策干预的能力，这是不符合市场原则的。这一章将追溯评级的历史，来说明这些数据是如何成为当代全球政治中的“全能武器”。

# 信用评级机构：从市场分析师到寡头垄断者

信用评级机构（CRA）并不总是像如今这样有影响力。为了理解它们的特权地位，我们有必要来看看一个多世纪以来全球金融结构的演变。

美国信用评级机构的历史开始于19世纪末20世纪初，当时美国公司数量和规模都在增长，尤其是那些建设长途铁路的公司。之前，工业发展总是由地方储蓄提供资金，但在其规模增大后，很快就需要有更多的资本投入，而这些资本无法仅仅由当地投资圈提供。人们需要挖掘其他更遥远的市场。然而，很显然的是这些新的投资者并没有第一手的投资项目的资料。在把钱投入新的项目之前，他们会要求对投资风险进行某种形式的第三方评估。通过这种评估，一方面有意购买债券的投资者会对项目有更大的确定性，而另一方面，公司也能够扩大潜在出资者的范围。正是由于对评估的需求不断增长，信用机构发现了有利可图的契机。金融分析家亨利·瓦纳姆·普尔（Henry Varnum Poor）是第一个试图系统地应对这种日益增长的需求的人，他对于正在进行中的工业发展类型做出了更准确的分析。1868年，他的《美国铁路手册》首次对这个行业进行了系统性的说明：“里程数、股票、债券、成本、收入、支出和组织架构；并描绘它们的兴起、发展和影响力”。

在信用评级机构兴起初期，它们评级的公司集中于铁路、轨道和车辆领域。它们会评估公司发行的债务证券，比如债券和有价证券，并根据债务偿还能力进行“评级”。在评级数据发行后，它们会被出售给投资者，从而使信用评级成为联系机构与投资者团体的一种纽带。

20世纪初，信用评级机构开始在美国金融体系中扮演重要角色。当时，财务分析师约翰·穆迪（John Moody）发表了他的第一份市场评估报告《穆迪工业及其他各类证券指南》，并创立了第一家评级机构穆迪公司（John Moody & Company）。几十年来，评级只不过是一种私人评估，与其他为投资者提供的任何咨询服务都没有什么不同。19世纪后期工业的蓬勃发展刺激了信贷的兴盛，信贷机构的力量不断增强，这也影响到了两次世界大战期间金融业，尤其是银行业的发展。传统上，银行一直是企业和公司的主要资金来源。它们是贷方（主要是私人储蓄，即存款）和借方之间的纽带，信用违约（及其风险）由银行

承担，而不影响存款人。因此，银行“是一种代表集体行为的，介于国家与市场之间的混合型机构”，控制风险，为“政治当局和借贷双方减少不确定性”。然而，经济大萧条时期和之后资本成本下降的压力增加，激励了银行（以金融和投资产品的形式）直接参与市场，最终降低了它们作为贷方和借方之间调解者的作用。由此，这种“非中介化”将银行的角色由自我调节的代理人转变为了市场的参与者。

（脱离银行调节后）资金分配的分散性直接与借贷双方相关联，这加剧了市场信息不对称的问题，从而也增强了评级机构的作用。

1936年，美国货币监理署强制要求银行必须持有至少具有中低风险的债券来将评级制度化。这个规定是为了防止银行用公共资金进行“投机级投资”，却还产生了间接效应，它通过赋予信用评级“法律效力”，使评级机构获得了银行债券持有者。共有四家机构拥有了这种准制度化的角色：标准统计局（Standard Statistics Bureau）、标准普尔公司、约翰·穆迪公司和惠誉出版公司（Fitch Publishing Company）。

1975年，负责监督美国金融市场的证券交易委员会（SEC）要求证券经纪人以高评级的资产来达到资本要求，使这些机构的定位得到了进一步巩固。由于证券交易委员会担心未经授权的机构可能会使评级市场膨胀，他们提出了“国家认可的证券评级机构”（NRSRO）的概念，并将其授予三大机构，也就是所谓的“三巨头”：穆迪公司、标准普尔和惠誉评级。从那时起，公共监管就只接受这三家公司的评级结果。其他金融监管机构也纷纷效仿公共监管，由此“这三家公司对债券安全性的判断就成为美国大多数主要金融机构债券投资组合的官方决定因素”。

从理论上来说，信用评级机构应该通过纠正贷款人和借款人之间的信息不对称来提高金融中介的效率。经济学家喜欢用委托人和代理人的关系来描述这个问题，也就是投资资金的人和负责投资的人都能收获成果。在这种观点下，评级机构被描述为“中立”机构，它们以一种客观和技术性的方式为市场参与者收集和提供信息。这种观点也反映了评级机构自身偏好的自我形象。例如，投资银行家团体——证券业和金融市场协会（SIFMA）表示，信用评级机构能“客观地分析一家公司或一种证券的信誉度”。

在信用评级机构出现的头几十年内，它们通过对发债公司评级和向投资者收取订阅费来赚钱。这使机构大体上对投资界比较负责和透明。这种状况在20世纪70年代发生了变化，主要是由于国际政治经济的一系列发展。随着布雷顿森林体

系的崩溃和随之而来的全球金融资本市场自由化，信贷评级的需求呈指数级增长，因为这有助于投资者评估和比较金融工具在全球市场的信誉和相关风险。因此，信用评级机构将其产品组合从公司评估扩大到对各种金融产品的评级，而这很快会成为它们的主要收入来源。与对实际存在的行业（通常被称为“实体经济”）的评估不同，对金融业的调研需要大量的抽象思维，这增加了评级的复杂性，使其方法不那么直观和对审查开放。与此同时，由于防止投资人分享评级资料（这会威胁到信用评级机构将其作为受版权保护的材料进行出售的能力）变得越来越困难，该机构的经营模式发生了根本性的转变，由订阅收费变为了评级收费。在这种新的安排下，被评级的公司开始为评级机构的评估付费，而投资人则能免费获得评估报告。

毋庸置疑，评级机构新的经营模式引发了各级利益冲突。由于企业需要为自身评级付费，信用评级机构开始有动力为企业进行夸大的评估，以满足和留住客户。从理论上说，这种发行人支付模式应该配以“安全装置”：由于信用评级机构在投资界的竞争主要看其声誉，它们应该避免夸大企业成绩，故意抬高企业的评级，否则会导致自身信誉降低。确实，在理论上是这样的。但实际上，在评级领域内呈现的是寡头垄断的权力分配，只有三家机构的支配型主导大大减少了行业竞争，反而鼓励了它们采用共同策略和企业联合。尽管美国现在有几家国家认可的证券评级机构来评估数以百万的债务，却只有三巨头有能力使市场和国家失去平衡。在这个行业缺乏竞争的同时，它的进入门槛也很高，所以寡头垄断会带来非常高的利润率。在金融危机之前的2005年至2007年，标准普尔的经营利润比2002年至2004年增长了73%。在2008年房地产泡沫破裂之前，人们对几十年来实行的评级制进行了调查，其中显示，有诸多因素都会导致评级的缺陷和可疑，包括“市场份额的驱动力、投资银行在提高评级上的压力、不准确的评级模型以及评级和监督资源的不足”。

## 数据背后：非正当的交易

有一个问题涉及了信用评级机构工作的核心：当投资人投资于债券或其他证券时，他们无法收回本金（以及承诺的利息）的风险是什么？

作为风险管理者，信用评级机构给出了明显比较复杂的系统，将“风险”转化为可度量的产出。他们会关注资产的特点和质量。他们会采访知情者、内部人士和其他所谓的专家。在进行初步评估后，他们会告知发行人暂定的评级，为后者提供重诉的机会。虽然评级量表在不同的评级机构之间有所不同，但它们都遵循基于字母数据符号的相似模型。三巨头都是用AAA来代表最优质的投资项目。BBB级别以上（穆迪公司为Baa级以上）的投资项目被视为“投资级”。评级越接近C，表示其越具投机性。对于标准普尔和惠誉来说，D就表示违约。尽管他们都采用字母符号形式，但这些评级的分数是由分析师来制定的，其具体方法并不可审查。在评级发布后，机构还会监督被评级的债券，对其进行可能的升级或降级，或将其置于“观察名单”中。由于规定要求发行人撰写详尽的招股说明书列出投资组合，信用评级机构的评级对许多投资人来说仅仅是“良好的例行公事审批印章”。简明的力量（以信用评级机构的分数为例）胜过财务分析的复杂性。市场使用评级作为投资指导，导致了在实际应用中，人们往往会认为只依靠信用评级就能作出买卖证券的决定。

信用评级机构的作用是引导和形成投资人的决策，从而建立一个引导市场行为的制度框架。就如同数据在治理中起到的作用一样，许多投资者会相信评级而“放弃自我判断”，他们对评级的权威性的信任“不是基于任何特定公告的优点，而是基于对权威本身的公正性的信任”。信用评级机构自身则避开了关于声誉的论争。比方说，标准普尔公司宣称“声誉比收益更重要”。穆迪公司曾经表示，“我们身处这样一个行业，在这里声誉资本更为重要”，而“驱动我们工作的主要因素是维系我们的过往成就。那才是我们的必需”。然而，一些批评者认为，“只有当信用评级机构的很大一部分收入不是来源于复杂产品的评级时，声誉论才起作用”，而如今的评级机构已经不再具备这种条件。其他一部分人则认为，由于评级不考虑一些其他的显著方面，如流动性和价格波动性，所以评级机构并不是旨在对投资决策进行全面指导。



还有一些人认为，评级机构只会关注那些已经发现问题债务人的市场。一些实证研究得出结论，“与市场信号相比，评级很少或没有信息价值增加”。当主权债务上升时，市场似乎会缓慢变化，有时也会迅速而轻率地产生变化。所以，评级制度仅仅依赖于“对可信者的信任”，其他人也会利用这些机构的评估来指导他们的决策，这就是为什么滚雪球效应如此频繁发生的原因。（本书分享更多索搜@雅书）

早在1985年，社会学家哈里森·怀特（Harrison White）就认为，由于金融业的专业化程度越来越高，存在着控制反转的趋势：“在代理人成为专门的供应商后，委托人就开始处于代理人的控制下。”在分散市场中，资本分配的选择因为假定是合理的、个体的，所以是独立的。核心问题是信用评级机构对这种选择进行干预。信用评级机构有权力改变资本的流向，这超越了“单一交易的个人主义认知行为”。因此，评级很容易成为自我应验的预言，因为评级机构大大影响了数以百万计的大小投资者的选择。此外，这种评级往往是顺周期性的，所以会人为地强化商业周期。由于评级机构的规模和市场支配力，它们可能会逾越信息提供者的角色，扭曲信贷市场。

使情况变复杂的是，机构声称其评级只是一种意见。在美国，受到起诉威胁的信用评级机构一再援引宪法第一修正案，由于评级受言论和表达自由条目的保护，对他们的指控由此被撤销。评级机构倾向于把自己比作出版公司或是财经新闻工作者，认为他们都仅仅是意见的发布者。因此，人们对于评级机构运作的态度似乎是自相矛盾的：监管标准以信用评级为基础，但对于评级如何制定，却不存在直接的监督。正如联合国贸易和发展会议（UNCTAD）的一份讨论文件强调的那样，评级机构“对于如何将相对权重分配给每个因素几乎没有提供任何指导”。在他们使用的一般标准和实际评级之间建立明确联系也是相当困难的。评级只是为投资提供“意见”，这种说法在很大程度上使评级机构不用顾虑传统投资建议中对透明度和问责的要求，这“能够使评级机构免于不正确或有误导性陈述的私人诉讼”。而且，“投资人从来没有投入大量资源以改进评级机构的运作状况，这可能是因为在此过程中缺少透明度。”

主要的信用评级机构通过为定向发行人做评估来赚取利润，这为他们提供了一个发布高质量评级的强大激励机制，因为高评级能够确保发行人在下次评估中还会与自己合作，而不是找另一家机构。投资银行为得到最好的评级“货比三

家”，已经是一种普遍的做法，它们甚至会“非正式地咨询两家评级机构，来制造一种竞争关系以获得比较高的评级”。从本质上来说，信用评级机构旨在实现两个主要的目标，这两个目标却往往是互相排斥的：获得最高的利润和客观评价客户，而对客户的评价又反过来决定了机构能获得的利润。这就会造成一种偏向，使机构的主要目标“客观性”更加难以实现。正如穆迪前高级分析师哈林顿（W.J. Harrington）所揭示的那样，高层管理者有着根本上的控制力。他们会对分析师说：“时间到了，我们来开会，都投赞成票。”而分析师提出的问题将会被驳回，或者直接搁置，他们会说“把那个记录下来”或者“很高兴你提出了这个问题”，但之后什么都不会发生。

由于业务中的支付结构，债务发行人（作为客户）会对机构（作为服务提供者）有非常大的影响。在2008年特许金融分析师协会（CFA Institute）对1956名投资专业人士进行的一项调查中，有11%的受访者表示，“他们已经发现有信用评级机构因为发行人、担保人或是投资人给予的压力，而改动了债券评级”。大约有一半的受访者认为，压力会表现为威胁的形式，比如告知机构“我们会把未来的评级业务都交给另一家”评级机构。CFA的调查进一步显示，许多受访者认为评级机构运作中的“支付结构”造成了最严重的利益冲突。显然，债务发行机构希望能获得最高的评级。当他们向评级机构支付评估费时，是他们占有控制地位。因此，在2001年安然公司申请破产之前，所有主要信用评级机构给予它最高评级也就不足为奇了。由于安然公司的高层管理人员担心评级降低将危及戴那基能源公司（Dynegy）对其接下来的拯救计划，所以要求评级机构给予它较好的评估结果，评级机构看上去也愿意合作。2004年，食品跨国公司帕玛拉特（Parmalat）的垮台显示，该公司先前在一些财务顾问的支持和评级机构的共谋下“做了假账”。该事件后，欧盟要求为评级机构制定更严格的规定。欧盟委员会和欧洲中央银行开始报告评级机构运作中的利益冲突，告诉人们他们如何建议发行人包装债券，以便授予AAA评级，也披露了评级机构未能提醒投资者注意次级抵押贷款市场危险的过失。然而，这些计划中的大部分都半途而废了，因为欧洲银行监管委员会认为它们是不必要的。

一些业内人士在评论大多数信用评级机构所采用的方法时已经证实，评估过程超过了简单的审查和鉴定，往往需要更多的“个人”扭转。对审查人员来说，与客户见面讨论保级甚至升级的选择，似乎是一种常见做法。正如芝加哥联邦住宅贷款银行（一家据说被穆迪和标准普尔都评为AAA评级的银行）前总裁回忆中的

那样，评级机构代表的来访是一种例常。“他们会说，‘这就是应该付的费用。’而我会说，‘这过分了。’他们重复，‘这就是应该付的费用。’最后，我会说，‘好的。’没有评级，你就无法出售债券。”一些信用评级机构也被指控勒索他们的客户，威胁他们若是要换评级机构，将立即对其进行降级处理。2007年，在美国参议院调查常设小组委员会听证会上，穆迪投资者服务公司一位匿名的常务董事反问道：

我们为什么不设想，信贷在宽松后会收紧，房价会在上涨后再次回落？毕竟大多数的经济事件都是周期性的，泡沫会不可避免地破裂。综合起来，这些错误使我们在信用分析方面看起来无能为力，或者就像是我们把自己的灵魂卖给了一个叫“收益”的魔鬼，也可能两者皆有之。

# 评级政策和全球金融危机

随着金融市场的多样化，信用评级机构所分析的产品范围不断扩大，从商业债券到整个公司，到主权债务，最后是过去几十年引入的无数新债务工具，其中最引人瞩目的是证券化，包括臭名昭著的担保债务凭证（CDO）和信用违约互换（CDS）。从20世纪90年代后期起，信用评估机构由此获得了丰厚的收益，因为它们会与金融公司合作制造以次级抵押贷款为基础的担保债务凭证。从那以后，标准普尔、穆迪和惠誉对这些证券进行评级的收费就达到了普通债券评级的三倍之多。

并且，随着金融市场一体化和全球化程度的提高，曾经的美国现象成为全球标准。自20世纪90年代以来，信用评级机构几乎渗透到了国际市场的各个地方。欧洲和亚洲的信用评级机构办事处如雨后春笋般涌现。如今，全世界的债务发行者，特别是工业化程度较低国家的债务发行者，都会认为从两个主要评级机构标准普尔和穆迪公司中的任何一个（或两个）获得评级，是一件必不可少的事情。在拉丁美洲，自20世纪80年代的债务危机后，债券市场的增长推动了对主权信用评级的需求。这些债券由发展中的经济体发行，有一定的风险，但收益率高，这就需要有这些国家金融信用的第三方信息支持。2002年4月，当时的美国国务卿科林·鲍威尔（Colin Powell）宣布，美国政府正在计划帮助非洲国家获得主权信用评级，因为这将有助于“鼓吹资本”以促进经济增长。就在一个月之后，联合国开发计划署（UNDP）与标准普尔开展了一项合作，支持在整个非洲推行系统性的主权信用评级。2003年，它还与纽约证券交易所以及非洲证券交易所协会合作举办了首届非洲资本市场发展论坛。最初的主权债券发行人（即国家）确实要通过谋求高评级来预估债券发行，但随着时间的推移，高评级的归属开始与各国的国际地位挂钩。这会帮助政府在全球面前呈现出透明且负责任的投资伙伴形象。通过评级，世界各国开始努力“获得来自国际资本市场的‘认可印章’”。

如今，信用评级机构不仅仅会从市场参与者的信任中获取影响力。各国也会通过将评级纳入财政条例来积极推进评级，并将评级制度化。一些评论员认为，评级机构“被普遍的市场惯例和公共监管赋予了参考地位”。通过转移自己在全球金融市场监管中的尽职调查责任，公共部门大部分参考了私人评级机构的判断来为中央银行确定贷款担保的资格，以及评估公共和主权财富基金的投资决策。早

在20世纪30年代，美国的养老基金只允许按照相关评级持有资产；此后，在需要对债务进行质量评估的任何时候，美国的金融监管机构都要诉诸信用评级机构。在美国经济中，信用评级被用来提高某些金融机构（如银行和保险公司）对投资限制的风险敏感度，为评级债券发行人确定差别披露要求，并调整对商业和投资银行的资本储备金要求。对于欧盟来说，1993年的《资本充足指引》促进了相似的趋势，该指引规定，企业必须为其持有的非投资级证券留出更多资本。具体而言，该指引规定，与在欧盟范围内交易的金融工具相关的违约风险必须由“主管当局认可的至少两家信用评级机构”进行评估，或当它们没有被其他机构评为投资级以下时，由“其中一家信用评级机构”评估。鉴于主要信用评级机构强烈的联盟倾向，评级总体上符合当局的要求并不令人意外。

《巴塞尔协议》进一步加强了评级机构的影响力。《巴塞尔协议》是由中央银行行长制定的一系列银行业规定，概述了对银行业的要求和建议。《巴塞尔协议II》（于2004年制定）特别规定了一项国际标准，以控制投资银行为了防范各种金融和运营风险所需要具有的资本量。这项标准制定的最初目的是使国际金融体系不受大型银行崩溃的连锁效应影响，但它最终因给予信用评级机构确定银行净资本储备金额度的权力，即银行必须为储备拨出多少资金来防止潜在损失，而造成了一个扭曲的控制和评价体系。

如今，债券发行人在法律上有义务寻求评级，以便出售债券并进入国际资本市场，这使得信用评级对他们来说非常有价值。这并不是因为评级中包含了许多信息，而是因为评级后他们会被授予各种监管许可。根据财经新闻媒体彭博社（Bloomberg）的报道，信用评级机构的覆盖范围“几乎延伸到了金融体系的每一个角落，从银行到管理机构的每一个人都已陷入评级体系中”。

利用信用评级机构对金融市场监管进行评估，有效地促使监管当局在提供公共利益信息上“退位”，它们将这种角色转交给了私人控股的寡头垄断企业。它们被政府赋予实际监管者的非正式地位，以避免自由市场经济学家将其视为“混乱监管”或“高昂成本监管”，但现在已经演化成了严重的公众过失案例。结果，国家和超国家权威的监管欠缺加剧了市场的持续衰落，这种情况如今又因银行体系的全球一体化而日益恶化。这套政策导致信用评级机构比原来的中介机构更加声称有利于信息交换、降低交易成本。由于各国“将许多监管职能外包给评级机构”，后者已然成为世界上最强大的市场守门人。

美国和欧洲都进行了一些改革的尝试，但几乎没有在监管和市场准入实践中产生任何变化。2006年，美国政府通过了《信用评级机构改革法案》，要求证券交易委员会负责该行业的监管。同年，欧盟重新修订了《资本充足指引》的规定，重新确立了信贷机构的作用，同时阐明了投资公司在高评级产品交易中的义务。此外，欧洲证券及市场管理局也受托开展调查并进行检查，作出相关罚款和运作禁令。在全世界范围内，信用评级也受到了国际证监会组织2004年颁布的《信用评级机构基础行为规范》的约束。当然，对该规范的依从仍然是自愿的，这保证了“信用评级机构的独立性，或是及时发布评级意见的能力”。

尽管政府当局为这个行业的改革作出了努力，信用评级机构的三巨头仍然保持着它们的一切权力和影响力，这一点在2008年金融危机爆发，以及随之而来的欧美主权债务危机中变得更加明显。由于信用评级机构在立法发展和国债“突然产生意外降级”后评级修改的滞后，它们会使市场失去平衡，从根本上影响世界主要经济体的公共财政。2011年，标准普尔采取了一项前所未有的举措，“将美国政府从无风险借款名单中除名”，这次降级引发了奥巴马政府的愤慨。美国政府对机构进行了攻击，认为“该公司犯了一个严重的数学错误”。美国政府还与证券交易委员会合作，对标准普尔“在房地产热潮期间对支持抵押贷款的证券的压倒性好评”开展了调查。调查人员还考察了公司分析师想要降低抵押贷款债券评级却被管理层否决的案例。对于证券交易委员会来说，评级机构“在为投资者的利益服务方面显然是失败的”。2013年，美国司法部起诉标准普尔涉嫌民事欺诈，“认为该公司抬高了抵押贷款投资的评级，并在金融危机来临时使其陷入崩溃”。

历史性的美国评级下调不是基于纯粹的财务评估，而是基于对美国政治策略的总体评估，这是金融分析师无法做到的。标准普尔的分析师在降低评级的理由中写道：

从更广泛的层次上来说，降级反映了我们的观点，即美国的政策制定和政治制度的有效性、稳定性和可预测性在当前的财政和经济挑战中已经有所减弱了。

因此，政府不仅给予了评级“法律效力”，还建立了它在国家政治策略上的发言权。在欧洲，尽管当局在2010年和2011年提出了几个救市计划，三巨头仍然将希腊、爱尔兰和葡萄牙等国家的债券评级降为了“垃圾级”。欧盟因此爆发了一场激烈的争论，由猛烈抨击信用评级机构的内部市场与服务委员米歇尔·巴尼耶

（Michel Barnier）主持，他们表示，“我们需要重建我们的政治主权，所以我们

不服从于市场主权”。德国前总理赫尔穆特·施密特（Helmut Schmidt）也坚持认为，“一些评级机构已经将在政治上应负责任的政府作为欧洲信誉的担保物”。

为了控制信用评级机构的权力，欧盟委员会提出了一些建议，包括要求金融公司每三年更换一次评级机构，分析师在四年后轮换客户，机构不能对包含其持股人经济利益的产品或机构进行评级。此外，主权降级的决定需在报告公开前至少24小时告知被评级的政府（而非此前的12个小时），而且只能在交易市场闭市后才能公布。由于人们认为反复降级会显示出“美国机构对欧洲债务的过分炒作”，有人甚至提议设立欧洲自己的信用评级机构。就连时任欧盟委员会主席若泽·曼努埃尔·巴罗佐（José Manuel Barroso）也对评级机构提出了质疑：

评级市场仅由三家机构垄断，却没有来自欧洲的机构，这一现象十分奇怪。这表明，在评估欧洲状况时，市场可能会存在一些偏见，毕竟欧洲人会比其他人更了解欧洲。

2011年10月，穆迪在二十多年来首次将意大利主权债务的评级下调，它认为，“意大利和欧盟其他地区经济增长的高度不确定性”仍然容易影响财政稳定。意大利政府在几周内就被迫屈服了。当时的总理西尔维奥·贝卢斯科尼（Silvio Berlusconi）在从政的17年间经受住了民间社会的抗议、罢工和一系列的关于腐败及性丑闻的司法指控，却不得不屈服于评级机构的权威，选择辞职。标准普尔于2012年1月13日（不久后被称为欧元区的“黑色星期五”）下调了9个欧元区国家的主权信用评级，包括剥夺法国和奥地利等国的AAA主权信用评级。欧洲的公共机构试图以多种理由起诉信用评级机构，其中包括对标准普尔前总裁德文·夏尔马（Deven Sharma）的指控。随着欧债危机加深，欧盟委员会关于评级监管的文件也表明，在财务危机状况中，信用评级机构不得发布主权债务的降级信息。

这种改变信用评级机构立场的企图遭到了业界的抵制。在美国，评级机构开始拒绝给债券发行人评级，这很快暂停了未评级债券的发行。在短期内，美国证券交易委员会也暂时放弃了评级要求。在欧洲，三巨头将欧元区的所有国家置于负面信贷观察名单中，这使投资者人心惶惶，更进一步加剧了欧债的紧张局势。此后，在2012年7月，穆迪瞄准了欧洲强国德国，出手将其评级展望由稳定下调到了负面。

政策制定者们终于意识到，由于政府和中央银行规定中的疏忽，信用评级机

构已经与银行和金融市场监管密不可分。如果连续的评级调整引发了各国各大洲采取二级市场冲销行动，各国政府便无法再采取行动。一方面，信用评级机构在国家和超国家的政策制定上不断加大影响，导致了政府合法干预范围的缩小。另一方面，他们稳定金融局势的方式与政治家的当务之急相冲突，导致在提出社会司法改革时出现不正当的财务惩罚机制，在实行财政紧缩措施时则被批准。评级决定加强了欧洲的社会困境，因为“政府借款利率上升意味着更多纳税人的钱会被支付给金融投资者，而不是花在公共服务和公共投资上”。因此，评级机构似乎不是在为解决主权债务问题做出贡献，而是在恶化世界各地的社会状况。

对于借款国家来说，评级下调会对它们获得信贷和借贷成本产生负面影响。此外，信用评级机构降低信用评分可能造成恶性循环，不仅使该国的利率上升，与金融机构的其他合约也可能受到不利影响，从而导致更多的开支和进一步的信用降低。有些人认为，评级机构能作为一个很好的机构“井蛙之见”的例子。通过“呈现‘确定性等价’<sup>[1]</sup>”，评级机构既能消减不确定性又能（再）制造不确定性”，而不是加强金融稳定性。

在金融和经济危机中，信用评级机构的行为本质上是保守的，也就是说，它们要么做出降级评估，要么保持之前的评级。评级下调，或存在下调的威胁都会削弱国家作为债务发行人的财务状况。评级降低是对国家债务风险提高的正式承认，风险高就要通过提高回报率来弥补。而较高的回报率不利于国库对现有债务存量进行再融资的能力、新债务的发行、通过财政政策对短期宏观经济冲击的应对，以及对旨在增加财政收入以偿还债务的短期投资的管理。由于税收收入是周期性的，信用评级机构认为它不够可靠，而倾向于采取私有化制度和紧缩措施，即公共资源私有化和减少支出。国家对市场的影响日益受限，于是引入市场机制，来提供更大比例的公共服务。

虽然主权信用评级不对国家收费，但是其中仍然存在不正当的激励措施。评级机构躲过了2007—2008年的经济危机、拉美和亚洲的债务危机、安然公司的违约和之前其他的欺诈性公司，如今对市场走势有些过于敏感，以至于评估也变得不稳定。此外，在经济繁荣时期的乐观评级和在泡沫破裂时期的突然降级引发了人们对舞弊行为的怀疑。当这些机构向购买评级服务的同一客户提供咨询服务时，问题尤其严重，例如向政府建议如何构建安全保障体系，然后根据这些建议作出评级。



全球金融危机已经证实，评级机构是有权势的政治行动者。它们不仅提供信息，还会在公司和其他公共机构做决定时，帮助构建起整个框架。评级机构并不像人们曾经期望的那样，是中立的、技术性的和客观的仲裁员。相反，它们首先会通过控制信息和调整评判，来组织、协调和“构建”资本市场。它们至少会拥有一个强大的“认知权威”，这可能和世界上其他的私人机构都不同。评级机构给出的数据影响了宏观经济政策，它们的评估会使整个国家和其国民陷入危机。

## 评级与非理性

信用评级机构对投资者来说非常重要，不是因为它们履行的信息职能，而是因为市场参与者相信信用评级机构给出的数据有重大意义，也就是，他们认为这些数据会影响到其他投资者的投资行为。公众对信用评级机构的期望在于其对一些资金流动和信誉“真相”的揭示，而这其实是错误地站在了对市场行为的理性认识上。当考察市场的运作方式时，相比实证分析，金融参与者显然会更容易受到他们对其他市场参与者反应期望的影响。凯恩斯在20世纪20年代后期首次提出的这种内生观点意味着，市场参与者集中于预测其他交易者可能做什么，以便从预见的行动和随后的市场变化中获得利润。市场通常对找到好的或者更好的投资并不感兴趣。它们对区分事实和谣言显得漠不关心。如果市场认为会有一定量的投资者相信某种谣言，这种谣言就足以变成事实，并会以此来转变投资计划，出售股票或债券。金融市场的社会和心理驱动因素，比方说一时的狂热、从众心理和集体非理性的其他特征，显然颠覆了完全自我调节的有效市场的假设，有助于解释日益频繁出现的非理性情况。这就是凯恩斯所说的“动物精神”，正是这种人类本能自然地引发了周期性的金融危机。诺贝尔奖获得者保罗·克鲁格曼（Paul Krugman）认为，信用评级机构的评级促进了市场的非理性，就其本身而论，为了达到市场监管的目的，信用评级应该被摒弃。他认为，政策制定者不应依赖市场机构进行自我调控，而应该建立可靠的规则和指导机制，以避免市场失灵。许多专家和顾问（包括衍生品行业的分析师）都支持取消评级的观点，认为这种做法能使“投资者和监管者戒除对评级如上瘾药物一般的依赖”。自由市场智库卡托研究所（Cato Institute）也主张政策制定者不应为监管目的而使用评级，以致承认了该行业“事实上的寡头垄断”。

联合国贸易和发展会议表示，“评级往往是有黏性的，它滞后于市场，且容易在市场真正发生变化时反应过度”，这会加剧金融危机，使其蔓延到整个国家。例如，在1997—1998年的亚洲金融风暴中，评级机构总在拖延而非主导事件，在关键阶段的过度反应加剧了繁荣与萧条的循环趋势，从而延长了危机周期。人们还研究了20世纪90年代中期墨西哥金融危机期间信用评级机构的作用。研究表明，大型机构对主权信用评级的负面公告显著提高了债券收益率和股市波动性，从而导致了不稳定的国际资本流动。正如我们看到的，降级通常会导致想购买政府债

券的投资人变少，这会转而导致政府债券更高的收益率。这种上涨之后会反映在每次国债拍卖的更高利率上，使得借贷成本更高，从而导致预算赤字增加，再造成进一步的降级，以此又开始了新的循环。

根据这些机构的说法，评级调整的滞后性是为了提供长期观点，而不是即时评估。这一策略意味着信用评级机构总是会有延迟，它们认为，任何特定的改革都不仅仅是可逆周期的初始部分，而是一段持续过程的开始。因此，比方说一个国家意图为改善财政状况做出努力，但可能会受到其偿债能力的持续消极前景（即使是错误的）的长期破坏。这种情况强化了这样一种观点，即信用评级机构呈现的数据并不能真正地提供任何信息增值，而只是单方面地影响投资者，遏制国家从负面冲击中恢复的能力，迫使它采取紧缩措施，这种措施对市场复苏的影响有限，同时还会给人口带来相当大的负担。

在过去，只有不到1%的AAA评级投资经历了彻底的违约。但在美国房地产泡沫破裂之后，“有大量AAA评级证券遭受了重大损失；一些证券公司甚至彻底破产了。”在2005年至2007年间，被标准普尔评为AAA级的担保债务凭证中，约有80%在美国房地产危机的动荡中被降至投资级以下。

美国参议院调查小组得出的结论是，过度乐观的评级支持了市场的猛涨，而后“突然的大幅下调.....会立刻引发金融危机”。经济合作与发展组织的首席经济学家宣布，信用评级机构“做出的判断会加快正在进行中的趋势”，他总结道，“这就像是推了一下已经站在悬崖边上的人。”

尽管公共监管机构已经明显认识到了信用评级机制的负面影响力，它们仍然会在制定遏制全球金融危机的政策时依赖于评级。美联储的定期资产抵押证券贷款工具为纳税人购买几万亿新证券提供了资金，这些新证券发行可以维持经济状况，而要获得这些贷款的条件就是资产抵押证券获得AAA评级。此外，美联储同意“当债务信用至少相当于A-1评级，即短期信贷的第二高评级时”，可以直接从商业票据发行公司购买商业票据，以此来为市场提供流动性。这些决定再次把信用评级机构置于了金融市场监管的中心，据估计，在2009年，三巨头可能会从纳税人的钱中获得高达4亿美元的收益。而在2008年经济崩溃期间，穆迪一家的营收为17.6亿美元，利润率为41%。

在欧洲，2009年的一项法规强制要求信用评级机构完成在欧洲证券监管委员

会的注册，才能在欧盟市场运营。主要的附加条款包括禁止评级机构提供咨询服务，若没有足够的高质量信息就禁止它们对金融工具进行评级，强制要求机构对其评级模型、方法和主要评级假设进行披露。正如彭博社指出的那样，公共机构已经采取了许多方法试图摆脱近年来最严重的衰退现象。然而，“没有人采取措施来对已经破碎的评级体系进行实质上的修复”。

彼得·费希尔（Peter Fischer）曾是美国最大的上市投资管理公司、位于纽约的贝莱德集团（Black Rock）的常务董事兼固定收益部的主管。他认为，评级机构完全应该被取代。在二十世纪初，股票被认为是复杂的，债券则被认为是简单的，为了告诉大家应该购买哪种债券而建立评级机构，看起来是有道理的，“但是，转到如今对信用的划分上，这就是一个非常奇怪的概念了”。在他看来，整个授权过程都是有缺陷的，因为它使市场和规则完全被少数公司掌控。评级机构变得就像是股票分析师，投资者直接为他们提出的意见而付费。

利用经济思维能完全理解风险的概念。事实上，风险是可以量化的：它可以被衡量，因此能被控制。能够控制风险的人便能预测未来，而能够预测未来的人会成为新的统治者。相反地，不确定意味着主观性。这处于猜测的领域，在指导决策时仅能运用常识和经验法则两个“参数”。不确定性可以被定义为纯粹的可能性而不能被数据化。没有传达权威意识的数据，就没有政治影响。在不确定性和主观性的领域里，每个人都是平等的。信奉自由市场的芝加哥学派的创始人弗兰克·奈特（Frank Knight）认为，不确定的条件使人们无法分配概率并预测未来进程。而对于奈特的对手之一约翰·凯恩斯（John Keynes）来说，不确定性消除了数据的力量，因为它阻止了对价格、战争或未来利率等的预测。尽管奈特和凯恩斯在政治经济学上持相反的观点，但他们都认同不确定性是经济活动中恒定的东西，它提供了人们获取利润的机会，同时也不可避免地让市场面临着亏损的可能性。这种开放式的企业经营，以不可预见的可能性为标志，评级机构试图通过借助隐蔽统计方法的复杂计算“驯服”它。它们宣称，将通过数学明显的客观性，把不确定性转化为风险。毫无疑问的是，计算技术确实取得了一定的进展，使它们能够扩大不确定性的边际，将更多的偶然事件转化为统计概率。

在经济学中，风险通常与最优均衡相关。它是一种设想的模型，能让你的成果在同行评审的期刊上发表，也能让你被数十亿美元的对冲基金公司聘用。不确定性只是未知的风险。就其本身而言，它们永远无法带你到达任何地方。风险管

理人员会受到尊重，给予报酬。而不确定性专家，就像是巫师一样，只存在于童话里。

评级过程是风险社会的核心。通过数据（或评级中的字母数据字符）显然的科学中立性，它传达了一种错误的控制和可预测性。穆迪也承认，“信用评级本质上是主观的”且“任何将信用评级简化为公式化方法的尝试都是误导性的，会导致严重的错误”。同样，标准普尔前总裁认为，评级过分地陷入了监管框架中。他恳请美国证券交易委员会在监管中避免提及评级公司，他觉得，“过度应用和不当应用都会带来非常大的风险”。

最后，可能有人会认为信用评级机构仅仅是社会用来处理难以忍受的不确定性的产品。信用评级机构的繁荣发展是因为它们能够非常狡猾地应用“经济和金融分析的客观化外衣”。它们隐匿于数据背后，“这样会比做出一个真正的困难判断要容易得多”。这些机构通过抓人眼球的计算来隐藏数据收集的基本概念，从而提出了一个使企业、政治家和投资者在传统上感觉非常舒适的现实观念。通过评级，信用评级机构实际上提供了一种虚假的自信心，它正对决策者的胃口，并能最终解释他们用所拥有的前所未有的权力来满足投资的渴望。然而，社会政治事件并不适合以数据概率的形式被解释。这就是为什么危机的出现能够揭开风险管理基本假设的不足，并让社会意识到竞争领域的不确定性，其难以靠风险管理来避免。

在政治表现上，信用评级机构的评级也曾一度因一种短视的叙述而兴起：过时和功能失调的市场将被胜利和有效的资本市场所取代。由评级架构起来的世界在金融危机的非理性状态下会迅速崩溃，在这个世界中，是数据政治而非公民政治，决定着如何治理我们的社会。最终，评级制度拥有了民主贫乏下的治理力量。

正如我们在前一章中谈到的那样，对数据的操纵不仅使治理过程有所偏向，同样也影响了测量结果和测量者之间的内在信任。在许多方面，评级机构都能完美地体现我们社会对于数据的信任和审计行业蓬勃发展的力量。在下一章中，我们将从金融世界转向气候变化中的政治和经济学。在这些领域，数据为气候科学家和所谓的怀疑论者之间的激烈争论提供了武器，最终，通过采用基于狭义的成本和收益概念的政策来回报市场力量。

## 注解：

[1] 在财务理论中，确定性等价是指在完美的和完全的资本市场中，有风险的未来收益可以转化为现值。

---

### 第三章

## 数据对全球变暖袖手旁观：气候变化的市场化

---

“公共政策本身可能沦为科技精英的俘虏。由联邦政府……形成项目统一分配，财力统一控制的局面，这种前景一直存在，而且应当给予重视。”

——德怀特·艾森豪威尔

虽然统计在政治上的运用和数据在全球治理中的力量是信用评级机构作用最有力的例证，但还有许多其他关键领域的衡量标准已经成为公共政策的中心话题。环境治理，特别是对气候变化议题的辩论，是过去四十年来“数据战争”全面展开的一个领域。不同的阵营应用数据、测量标尺、模型和指标来达到不同的目的。这个领域也变得如同蓬勃发展的评级行业一般，评审者占据了中心舞台，能够赚取大量金钱。

这四十年里，人们对世界整体环境状况的担忧日益增加。1972年联合国人类环境会议（通常被称为斯德哥尔摩会议）开启了一系列高级别首脑会议，世界环境与发展委员会，更广泛地被称为布伦特兰委员会，于1987年发表了第一份可持续发展报告。1988年，世界气象组织和联合国环境规划署共同建立了政府间气候变化专业委员会（IPCC）。随着该委员会的成立，来自世界各地的科学家们开始着手整理信息，以便在国家层面上达成对地球气候问题的共识。他们的主要目的是通过核查大量数据来说明地球究竟有没有变暖。政治家们等待着统计数据以便采取进一步行动，所以毫无疑问，科学家们觉得要迫切证明全球变暖确实正在发生。与此同时，许多智库和持相反意见的科学家开始进行相反的研究，旨在揭露IPCC所做的工作；各类经济学家引入各种模型来评估气候变化引发改革的利弊，这引发了社会科学界的辩论、争执和意味深长的对比。其中，数据之战使主张行动者与论证不行动的经济优势的人形成了深刻分歧。环保组织、社会运动和民间社会的各种声音开始质疑工业革命以来发达经济体所追求的发展模式的基础，这使这种模式变得危在旦夕。

这些年来，数字政治封锁了市场方法与环境之间的密切联系，最终我们将间

接地接受气候怀疑论者和化石燃料工业中提出的建议。为了确定市场利益和自然利益之间的可接受的均衡，人们在气候变化治理中引入了狭义的经济推理以及其中一些传统的方法，特别是成本效益分析。数据模型的表面中立性最终导致了气候争论的市场化，其中可持续发展、碳排放与交易，以及绿色增长等概念成为一种语言手段，用来加强市场对所谓向低碳社会过渡的控制。



## 环境怀疑论和成本效益分析的兴起

历史学家内奥米·奥利斯克斯和埃里克·M·康韦在《贩卖怀疑的商人》一书中特别以美国为例，详细分析了一些工业游说团体、保守派“智库”、私人基金会和所谓的环境怀疑论者之间的联系。他们在书中呈现了这个“否认行业”的触手可及，以及它对美国政治制度的影响。自20世纪70年代以来，这个行业的策略就是在达成科学共识的领域（著名的例子是对癌症和吸烟之间关系的研究）引发争议并制造疑虑。他们“反科学”的总体目标是反对政府在各个领域的监管，保障稳定的工业利益，尤其是在矿物燃料和军工领域。比如，他们曾为里根总统的战略防御倡议（亦称“星球大战计划”）提供科学报道支持。在烟草研究所的支持下，这一战略在过去的几十年中一直被用来阻碍改革发展，其中涉及了酸雨、臭氧层空洞和气候变化等领域。这种观点的主要推动者是强势的科学家，比如在“二战”期间参与曼哈顿计划的物理学家威廉·尼伦贝格（William Nierenberg），美国国家科学院前主席、北大西洋公约组织顾问弗雷德里克·塞茨（Fredrick Seitz），美国海军科学家西格弗里德·弗雷德·辛格（Siegfried Fred Singer）和自由主义卡托研究所的研究员、《消融：科学家、政客及媒体对全球变暖的失实预测》一书作者、气候学家帕特里克·迈克尔斯（Patrick Michaels）。他们都以不同方式与化石燃料行业游说团体或军工企业联系在一起，并在里根政府（1981—1989年）时期和乔治·H·布什任期（1989—1993年）内，在通报环境政策制定方面发挥了关键作用。尼伦贝格和塞茨共同创办了乔治·C·马歇尔研究所，这是一个强大的保守派智库，成立于1984年，旨在支持里根的核防御计划，反对对忧思科学家联盟等组织的批评。在接下来的几年中，研究所开始资助“替代性”研究，对吸烟（包括被动吸烟）的致癌性质、产生酸雨的原因和后果以及臭氧层的损耗进行争论。弗雷德·辛格曾为埃克森美孚和壳牌石油公司以及洛克希德马丁等军工企业提供咨询服务。之后，他成为弗吉尼亚大学教授，组建了科学与环境政策项目，这是一个就环境问题的主流科学共识进行争论的倡导性团体。帕特里克·迈克尔斯在接受美国有线电视新闻网（CNN）主持人法里德·扎卡里亚（Fareed Zakaria）的采访时承认，他的收入和研究经费中大约有“40%”来自石油行业。这一说法引起了美国政策制定者的指责，他们认为迈克尔斯在2009年美国国会能源和商业委员会举行之前的证词“误导”了国会。

在过去的几十年里，大量的基金会和智库一直在积极支持美国的环境怀疑论者，他们的争议立场被自满的媒体进一步放大，其中一些媒体故意赞扬了非正统的观点，另一些则简单地采用了所谓的“一视同仁原则”，这是一种建立于1949年电视行业崛起时的准则，要求广播记者在对公众关注的争议性问题报道中，给予争论双方同等的播出时间。社会学家赖利·邓拉普（Riley Dunlap）和艾伦·麦克里特（Aaron McCright）撰写了2011年《牛津气候变化与社会手册》的一个主题单元，他们认为，“否认气候变化的组织”围绕着少数几个商业集团（包括化石燃料巨头埃克森美孚公司、美国石油学会、西方燃料协会，以及自然资源组织，如美国国家矿业协会和美国林业及纸业协会），这些组织为乔治·C·马歇尔研究所、美国企业研究所、卡托研究所、竞争企业协会以及传统基金会等智库提供了系统资金。他们在过去几十年的游说和研究工作成功地为气候变化及其原因和影响的科学共识制造了一种“怀疑”的气氛。在少数情况下，他们也成功发起了比较流行的运动，比如已经解散的全球气候联盟（一个反对温室气体减少的商业运动）和美国清洁煤电联盟。

虽然怀疑论者的观点已经被科学界无数次地反驳过，但他们最重要的（也是最持久的）成功就是对功利主义推理的普及，这是处理环境治理决策的最好方式。他们与里根政府的亲密关系使他们不仅能够影响总统对于环境问题的观点，还能影响他解决问题的想法，在做出商业利益与生态问题之间任何潜在的权衡时，总统会倾向于采取以市场为基础的方法。这些学者有助于使决策者（如今仍然拥有权势）达成广泛的共识，他们容易达成这样一种观点，即环境保护（和保存）不应被视为促进人类福祉的基本价值观，而应看做经济发展的障碍。他们还认为，可持续性的三个要素（经济、社会、环境）是无法相互补充的，三者只能通过不断取舍，使人们基于效用的标准来运用和指导决策。例如，弗雷德·辛格本人，就是第一批支持在处理环境问题时使用成本效益分析的人。成本效益分析是一种对成本和收益进行的比较性衡量，通常与特定的决策、项目或政府政策相关，如今已经成为大多数国家环境政策事前评估中的普遍做法。要使成本和收益能够被比较，它们必须以货币形式表示，并且要调整到特定的时间范围。由于成本和收益经常出现在不同的时间点，特别是在环境保护等领域（这一代人承担成本，未来才能享受大部分收益），经济学家只能将贴现率应用于未来的收益中：理论上，这样做能消除时间差，并估计出整体收益是否超过了总体成本（即所谓的“净现值”）。当成本大于收益（以市场价格来衡量）时，经济合理性就要求人们不采取行动：不做任何事情才会有经济意义。美国米特里公司是美国军事装备

的主导力量，在1979年由其委托制定的空气污染控制成本和效益的报告中，辛格正是站在了这样的立场上，他赞成“保守的空气污染控制方法”，并提出了替代方案来“降低国家成本”。1982年，他因在环境评估领域应用成本效益方法做出贡献而获奖，并被邀请加入里根总统的酸雨同行评审专家组，该小组由他的朋友、同样是怀疑论者的威廉·尼伦贝格担任主席。

成本效益分析存在许多概念性的和方法上的问题，这些问题常常会导致在运用分析时，产生截然不同的结果和无尽的争议。我们可以用缓解全球变暖问题中的成本效益分析来举例。根据伦敦政治经济学院教授尼古拉斯·斯特恩（Nicholas Stern）2006年为英国政府所做的气候变化经济学的著名评论文章，强有力的早期行动对气候变化的好处远远大于不采取行动的代价。报告中指出，气候变化每年将造成国内生产总值5%至10%的损失，而采取行动避免大部分影响的成本只占全球收入的1%。然而，其他经济学家，包括诺贝尔奖获得者威廉·诺德豪斯

（William Nordhaus）等做出的另一种计算，则得到了相反的结果。卡托研究所进行的一项研究把门槛抬得更高，该研究表明，只有当气候变化每年对国内生产总值造成至少10%的损耗时，减排才是值得的。耶鲁大学教授罗伯特·门德尔松

（Robert Mendelsohn）遵循诺德豪斯的理论，认为社会应该将边际减排成本与边际损失平衡起来，他认为，“现在承担的代价比未来产生的代价更加沉重”，由此得出结论，当代人只应投资于部分减缓气候变化政策，这部分政策需要能获得“与无数市场部门替代方案中的竞争性投资相同的回报率”。有趣的是，在斯特恩2008年4月发表的一篇论文中，他通过指出环境怀疑论者所使用的传统武器来反驳他的批评者，即“不确定性和将来解决某些问题的前景，常常被作为拖延行动的理由”。他补充道，2007年IPCC“第三次评估报告”的结果证明，他此前宣称的尽快进行大胆改革是正确的。他还强调，尽管经济学家似乎愿意承担高风险，但将后果落于子孙后代的做法，“大多数人都会觉得不道德”。他承认他之前的评论文章可能有些错误，但从谨慎的想法考虑：“我们低估了风险……低估了气温上升带来的破坏……以及温度升高的可能性”。



成本效益分析这个概念所依据的贴现率，最终是由研究人员做出的个人价值判断。一般来说，经济学家喜欢把未来产生的收益赋予一个较低的价值，因为他们认为那时候的收入会更高，这考虑到了边际效用递减的原则，使后代更容易承

担成本较低的消费。他们还认为，技术进步将会产生更有效的方法来解决环境问题。此外，他们的模式优先考虑生活在当下的人，因此会排斥代际连带和长期可持续性价值观。正如伯克利经济学家、美国财政部前助理副部长布拉德·德朗（Brad Delong）所说的那样，我们的“推理存在根本性的缺陷”，因为我们总是“更急于现在和不久之后能获取的价值，远胜于遥远的未来”，并认为“一鸟在手，胜于二鸟在林”。事实上，无论成本效益模型看起来多么“精致”，事实仍然是，没有人能知道我们的后代将如何评判我们如今所做决定的价值，或者换句话说，我们如何估量因为现在承担了成本而给未来的子孙带来多少收益。

由于所有这些假设，辛格在酸雨同行评审专家组1984年的报告中得出结论，要量化控制空气污染的成本和收益是非常困难的（他实际上采用了声名狼藉的粗略估计“用十亿美元的方案来解决一百万美元的问题”，却没有提供任何有关他如何计算出这些数据的信息）。而在1979年，白宫环境质量委员会已经估计出，清洁空气的经济效益约为每年214亿美元。辛格建议采取基于市场的可转让排放额度制度，政府只需要设置最高限额（如今我们称之为“上限”），然后将这些权利出售给公司，公司就会通过使用或交易这项权利来取得经济收益。这正是布什政府在最后采取的方法。根据辛格的建议，美国政府于1990年启动了世界上第一个大规模交易计划，旨在遏制形成酸雨的气体二氧化硫的排放。2003年，美国环境保护署量化了过去十年酸雨预防措施的总体成本，每年约为88亿美元，而估计的收益则是该数据的十倍以上（每年1010亿美元到1090亿美元），从而驳斥了辛格的早期分析。尽管辛格的方法有众多缺陷，它最终还是会在环境治理理念的斗争中取得胜利。事实上，自《京都议定书》通过以来，排放交易计划的迅速扩散证实了怀疑论者主要论点的全球吸引力：无论如何都要避免监管，让市场进行统治。

## “气候门”事件：对气候数据的扭曲

从《京都议定书》通过到全球经济危机爆发的这些年，气候变化在国际政治中始终处于中心地位。尽管怀疑论者试图说服大众，科学并不具有决定性，全球各国政府和政治领导人仍为此召开了一系列高级别峰会，并不断作出相关承诺。2007年，IPCC与美国前副总统阿尔·戈尔（Al Gore）共同获得诺贝尔和平奖，他当时是积极改善全球变暖危机的行动者之一。由环保联盟发起全球运动激励了各大洲的数百万人。奥巴马的成功当选，加之国会两院中的绝大多数民主党人，似乎在向全世界保证，在美国近二十年反对任何形式的国际合作后，终于准备好承诺缔结一项具有约束力的国际条约。与此同时，世界各国正在对一项新的条约进行谈判，以取代《京都议定书》，所有的目光都集中在准备2009年12月将在哥本哈根召开的联合国气候变化大会。然后，就发生了“气候门”事件，它为气候科学家证明地球变暖提出的一些数据是否成立蒙上了一层阴影。由于大多数政策和经济分析（包括“斯特恩报告”）都依赖于气候科学家提供的数据，这无疑给怀疑论者提供了一个极佳的途径来抹黑所有预测的来源并引起公众注意。

在菲尔·琼斯（Phil Jones）学术生涯的大部分时间中，他都是一位高产但无名的气候科学家。自1998年来，他一直担任英国东安格利亚大学气候研究中心主任。琼斯与英国国家气象局密切合作，从全球数以千计的气象站收集信息。他负责所谓的仪器温度记录，也就是几百年来全球陆地表面和海洋温度波动的时间序列。2001年，他对地球温度的研究在IPCC第三次评估报告中起到了重要作用，之后在2007年，他的工作对第四次评估报告的结论产生了深刻的影响。该报告指出，“气候系统的变暖是非常明确的”，以及“自20世纪中叶以来，全球平均气温升高的大部分原因来自于由人类活动引起的温室气体浓度增大”。

几乎是一夜之间，琼斯本来平静的生活被一场全球性的丑闻压垮了。2009年11月，气候研究小组的服务器遭到黑客攻击，数千封电子邮件和文件被盗。被黑客窃取的信息被发送到了一些气候怀疑论者的网站，其中包括“气候监测”网（Climate Audit）。它可以说是气候领域最著名的持不同观点的博客之一，由加拿大数学家、前矿业顾问史蒂夫·麦金太尔（Steve McIntyre）建立。麦金太尔与竞争企业协会——一个长期以来对官方全球温度数据的有效性提出质疑的智库——有密切联系。随后，琼斯的私人通信被仔细检查，并在互联网上广泛传播。

在数以千计的网站和社交网络上都可以看到文件内容，最后，媒体宣布了这一消息。

不夸张地说，琼斯的电子邮件显得十分尴尬。在一些邮件中，他经常要求作者和同事避免把他们的数据公开，以免被对手利用。这显然违反了科学的开放性和英国的信息自由法，而设置该法律本来是为了给热爱科学的人提供研究机会。琼斯为麦金太尔和其合作者，经济学家罗斯·麦基特里克（Ross McKittrick）多年来对他数据的关注而感到痛心，“如果他们知道现在英国有信息自由法，我想我会删除这些文件，不会发给任何人”。其他信息还显示了气候科学家的合谋，他们“向发表质疑气候科学共识文章的期刊编辑施压”，会在IPCC官方报告中剔除一些更重要的分析，“即使我们因此必须重新定义同行评审”。琼斯的副手基思·布里法（Keith Briffa）在一封电子邮件中承认，他“一直在努力平衡科学与IPCC的需求，而这些需求并不总是相同的”。

当皇家气象学会要求所有期刊作者公布数据时，他们对其展开了投诉。琼斯威胁道，不会再向“任何皇家气象学会的期刊投递论文”。在2009年1月3日的一封电子邮件中，气候研究所的迈克·麦克拉肯（Mike McCracken）提出担忧，认为有关地球温度升高的预测可能是错误的。他建议考虑提出一个备用计划，比方说，提出硫酸盐会导致全球降温，“否则，怀疑论者的言论将压倒我们——世界真的正在变冷，我们的模型一点儿都不好，等等。而美国正准备严肃对待这个问题”。

从负责更新研究小组数据集的研究员、程序员伊恩·哈里斯（Ian Harris）保存的期刊中检索到的内容更让人吃惊。哈里斯在所谓的“自述文件”中指出，有些弱化统计相关性的案例存在缺失：“这里究竟发生了什么？好吧，没有‘假设’，我可以自己弥补。简直完美”。他还哀叹了“数据库令人绝望的状态”，其中“没有统一的数据完整性，它只是一个人们一发现了论点就往里添加的论点目录”。他总结道：

你无法想象这会让我付出多大的代价——事实上，我是在让运营者分配虚假的世界气象组织（WMO）代码！！但在这种情况下还会发生什么呢？他们会特别处理可疑出处的“主”数据库（呃，他们都是，而且会一直这样做）。

然而，让步意图最明显的信息，是由琼斯自己发出的。迈克尔·曼（Michael Mann）、雷蒙德·布拉德利（Raymond Bradley）和马尔科姆·休斯（Malcolm

Hughes）三人于1998年在《自然》杂志上发表了一篇分析全球温度模式的论文。在琼斯与三人的交谈中，他自夸了对一些方法性“技巧”的使用，以掩盖前五十年温度下降。琼斯、曼和同事们都认为，这些陈述脱离了上下文背景，在很大程度上被曲解了。然而，琼斯的坦白引起了人们对气候变化辩论中最受欢迎的标志的怀疑，即所谓的“曲棍球杆曲线”。由曼和他的同事们研究得出的“曲棍球杆曲线”显示了过去六个世纪地球的平均气温，他们通过仪器收集数据并绘制图线，反对基于古树年轮宽度产生的图像重构，以代表过去的温度变化史。根据他们的计算，在19世纪末之前，全球气温几乎没有变化。然后，在20世纪，气温数据开始上升，这也导致了图线的急剧上升，就像曲棍球球棍面一般。对于许多人，不仅仅是科学界人士来说，这张曲线图都能非常直观地表明，自从工业革命以来，人类确实已经改变了曾经稳定的气候模式。

这正是气候怀疑论者一直在等待的。在他们看来，这证明了气候学家一直在修补他们的数据，以显示地球气温的上升，众多批评者对此结果展开了大量争论。其中就包括怀疑论者大本营乔治·C·马歇尔研究所的创始主席，颇具影响力的弗雷德里克·塞茨。他在1996年的《华尔街日报》社论中严厉批评了IPCC的报告，并断言，“我在美国科学界的六十多年中，从未目睹过同行评审出现如此令人恐慌的舞弊现象”。怀疑论者认为，所谓的“黑客”其实是一位内部人士，他对研究小组采用的方法感到不满，决定揭发这一切。麦金太尔和麦基特里克写了大量论文质疑曲棍球杆曲线和“20世纪地球历经了惊人高温”结论的正确性，因为“气候门”事件的发生，他们更觉得自己的质疑是有据可循的。

几天之后，琼斯的邮箱遭到了辱骂性邮件的猛攻，其中有一些甚至威胁了他和他家人的生活。在窘迫和烦扰中，琼斯辞去了研究中心的工作。在接受《星期日泰晤士报》的采访时，他承认，自己曾考虑过自杀。调查小组启动了几次官方调查。其中最重要的是由英国下议院科学与技术委员会进行的，该调查得出的结论为，对琼斯及其团队的指控是不成立的。这个科学家小组采取了一种容易引起争议的交流方式，但调查组并没有发现任何证据表明科学家遗漏或是篡改了数据。2010年，进一步的调查发现，琼斯及其合作者们在“作为科学家的严谨和诚实”上并没有错误，因此琼斯得以在其学术岗位上复职，只是变更了职称。然而，气候研究小组确实没有达到人们对科学家普遍期望的开放精神，这主要是因为他们对信息和数据的共享感到抗拒。迈克尔·曼，作为宾夕法尼亚州立大学地球系统科学中心主任，也受到了校方的调查。审查其案件的小组得出的结论是，没有证

据证明数据是伪造的或遭到了破坏，曼也没有滥用特权和机密信息。然而，科学家们是否“偏离了学术界的公认做法”仍然有待考量。

尽管调查组并没有发现研究人员渎职的证据，尽管研究人员已经按计划恢复了工作，最重要的是，尽管国际社会已经肯定了他们对全球变暖研究贡献中的科学有效性，但是这次风波还是造成了对气候学术界的破坏。由于媒体已经见风使舵地转为“气候变化抨击者”，本来的一个小的公共关系问题，成为诽谤抹黑数十年的科学研究的历史性机会。例如，IPCC被指责在第四次报告中隐去了关键章节。尽管IPCC对此进行了否认，但其发表的文章中确实包含了一些错误，这也使得时任联合国秘书长潘基文要求对其进行独立审查。鉴于美国科学家参与了可疑的电子邮件通信，美国国会的共和党议员要求由司法部进行刑事调查。他们认为，该事件涉及的气候学家一直在“操纵数据，故意使用有缺陷的气候模式达成先入为主的结论”，因此，IPCC的共识“人类生产活动的排放不可逆转地导致了环境灾难”也同样被连累了。他们还呼吁美国环境保护署重新审视其与温室气体相关的风险立场，认为所谓的危害性发现，即工业和汽车排放会威胁到当代和后代的公共健康和福利，“应该被否认”。

正如《自然》杂志评论的那样，“（该事件）对气候科学的声誉造成了巨大的损害，对整个科学体系来说也是如此”。被媒体广泛采用来形容这一事件的“气候门”这个名字，无疑有助于创造一种怀疑和操纵的气氛，就像是气候科学家一直在黑暗的高校环境和实验室中策划一个恐怖的计划，以控制环境治理政策的制定。和预想的一样，阴谋论确实大量存在。科学与公共政策研究所（SPPI）是一个反对“为微不足道的利益作出大量经济或政治牺牲”的智库，其发布的一份报告被大量下载，报告中指出，气候小组“曾图谋对什么是同行评审科学进行重新定义”。报告认为这些“气候犯罪分子”应该“因为他们欺骗性地篡改科学数据，并压制与他们观点不相符的结果而被囚禁”，还总结道，“应该把他们制造出的不成问题的‘全球变暖’议题立刻搁置，今后不应再为此提出任何公共政策措施”。这份报告的作者，英国评论员克里斯托弗·蒙克顿（Christopher Monckton）在20世纪80年代曾经是首相玛格丽特·撒切尔（Margaret Thatcher）的科学顾问，他曾因为对艾滋病的看法而臭名昭著，他建议“定期对全体人口进行筛查，隔离所有病毒携带者”以避免艾滋病传播。在看似对主流科学研究结果的协同攻击中，SPPI发起了一系列诽谤运动，包括公开呼吁打压政府间气候变化专业委员会和逮捕阿尔·戈尔。在另一家怀疑论者机构——二氧化碳和全球变化研究中心的支持下，他们从其宣称的40



多个国家，600多家机构的1000多名科学家那里搜集了证据，以显示IPCC关于全球气温变化趋势的共识确实如他们的认识一般，是错误的。他们发行的书籍也由此被命名为《大气中二氧化碳浓度升高的诸多益处》。

对于怀疑论者和他们所代表的游说者来说，“气候门”的时机实在不能更有利了。就在黑客入侵几个星期后，世界各国领导人在哥本哈根会面，共同商议《京都议定书》的后续方案。环境保护主义者和关注环境问题的公民对第17次缔约方会议（通常被称为COP17）非常失望。不仅国际社会拒绝了达成共同长期协议的可能性，而且从那以后，所有国际峰会都被当作推迟解决环境问题的“昂贵”机会。美国这个环境怀疑主义的大本营也看到了公众对气候变化关注度的急剧下降。在2008年，有71%的美国公民认为全球变暖正在发生。但是，到2010年，这个数据已经下降到57%（不相信气候变化的公民比例从2008年的10%提高到了20%，而对此表示不确定的人增加到了23%）。然而，在那些相信气候变化的人当中，只有59%的人是“非常”或者“极其确定”气候变化正在发生的，这比2008年下降了13%。同样，只有约半数美国人对气候变化“感到担忧”，而在2008年，持有相同观点的人有63%。

哥本哈根共识中心（CCC）是从“气候门”争议中获利最大的智库之一。它在2002年，由“持怀疑态度的环保主义者”比约恩·隆伯格（Bjørn Lomborg）创立，包含了一小部分的研究人员和广泛的合作者网络，其中包括数名诺贝尔经济学家。哥本哈根共识中心的使命是“增长知识，在既定的问题中给出研究和真相的概述，这意味着优化措施始终基于证据”。“证据”这一提法当然非常有力，但他们却没有进一步说明，这些证据是由什么组成的。该中心的目的是回避意识形态和政治议程，完全依靠数据和经济推理。通过经常使用成本效益分析，他们将自己塑造为解决全球问题的新前沿。2008年，通过由五位知名经济学家〔包括自由市场经济主义者贾格迪什·巴格沃蒂（Jagdish Baghwati）和奥地利经济学家弗农·史密斯（Vernon Smith）等〕组成的小组，他们列出了应对气候变化最佳和最坏的方式。结果正如预期，最好的方式大都为非监管性方法，比如技术创新和气候工程，其中包括了碳封存。他们还建议更广泛地推行从工业化国家到工业化程度较低国家的技术转让，并为适应气候变化提供资金——支持那些旨在帮社会做好准备应对恶劣气候的项目，而不是支持那些试图促进变化减缓的法律提案，比方说在富裕国家引入烟尘排放税来应对气候债务。在2009年发表的一篇论文中，哥本哈根共识中心抛出了一系列数据来表达观点。其中估计，到2100年，气候变化导

致的累积福利损失大概会是收益的几个百分点：“也就是说，一个世纪的气候变化产生的损失约等同于一到两年的经济增长值”。那么为何要自寻烦恼呢？他们的结论是，气候变化不应该成为决策者优先考虑的问题，至少目前是这样。只有当社会成本更容易被接受的时候，才需要提出处理环境问题的具体措施。

哥本哈根共识中心的创始人认为，该组织通过倡导“对环境处理的经济手段”来加强对气候变化的循证思考。因此，他们会优先考虑资源货币化的解决方案。例如，当要在保护生物多样性的同时解决粮食短缺问题时，他们会强调通过研究和开发提高农业产量的重要性，“这样就可以用较少的土地提供更多粮食”。他们估计，若每年能为研究投入145亿美元，就有可能使作物的年增长率提高20%，畜牧业的增长率提高40%，从而在未来四十年内减少对自然的压力，保护生物多样性。总的来说，其宣称的获益将达到530亿美元：“每消费1美元，我们就会为生物多样性和气候带来7美元左右的利益”。

2012年，哥本哈根共识中心发表新“共识”，概述了未来几年最重要的全球性挑战。他们将气候变化问题排在了靠后的位置，并建议只用花费少量的公共资金（大约10亿美元）来开发气候工程的新领域，如平流层气溶胶注入（造成酸雨的前体物质二氧化硫会被连续注入平流层，形成一层气溶胶以反射阳光）和海洋云增亮技术（将海水混合到大气中，使云层更亮更具反光性）。他们还声称，气候变化对经济产生的影响被严重夸大了。在他们看来，最负面的影响可能会表现于农业和旅游业，“到21世纪中叶，平均每个国家的损失可能会达到GDP的0.5%左右”。但他们也指出，能较好地适应环境变化的人就能避免这些损害。农民会选择种植更适应高温的作物。人们会设计出新型的房子应对气温上升：“考虑到适应性，富裕的国家会适应全球变暖的负面影响，并积极利用气候变化，在全球变暖的积极效应中产生约等于GDP半个百分点的价值”。他们再次表明，全球变暖没有什么大不了。恰恰相反，气候变化可能会为所有人带来经济福祉。

经济推理的应用和其宣称的中立原则，可以变得非常吸引人。事实上，对成本效益分析的依赖从根本上说是一个可怕的现象，它过分简化了社会问题的多维性，使权力结构始终反对解决长期存在的全球性问题，而我们对这种持续状态视而不见。在下一部分中，我们一起来看看这是为什么。

# 气候市场

怀疑论者对气候科学家提出的主要指责之一是，他们对地球温度状况做出的结论产生了由“绿色经济”投资者、碳交易市场和抵消机制组成的牟利行业。气候怀疑论者的报告以相当直接和咄咄逼人的方式指出了这个问题：

所有从碳交易或“绿色投资”或联合国任何对气候的无效投资中赚钱的“全球变暖”牟利者都应受到明确的警告。其盈利活动的基础是虚无的、欺诈性的，他们将被控告、起诉，因欺骗而入狱，他们赚的钱将以洗钱之名被没收。

在绿色产业中，把气候学家与投机者联系在一起至少是不现实的（在某些方面，可以说是具有冒犯性的）。气候学家强调了一个问题（也就是地球温度升高和温室气体浓度），但并不站在某种立场上给出解决问题的最佳政策。事实上，“绿色增长”范式是由企业和决策者（及其经济顾问）发明的，而非气候学家。这个领域的大部分投资实际上来自长期受益于气候变化否定主义和管制放松的高污染企业（如石油公司和采掘工业）。石油和煤炭行业在很大程度上控制了全球的碳市场，同时也致力于大部分的气候工程投资（如碳捕获和储存）。对于其中许多企业来说，气候变化已经成为一项利润丰厚的业务。正如我们看到的那样，气候怀疑论者弗雷德·辛格就是环境问题市场解决方案的有力支持者。他所支持的“酸雨计划”成为历史上第一个大规模的排放交易体系，世界其他国家将以此为例来制定由市场驱动的缓解气候变化政策。目前碳市场已经出现在各大洲，包括欧盟、北美（如区域温室气体减排行动和美国最大的加州碳排放交易机制）、新西兰、澳大利亚、日本（东京）和中国。

排放交易计划的设计中包含了多种机制。最普遍的区别在“限额与交易”和“基线与信用”之间。在前一种情况下，政府当局会设定特定的排放上限（如通过测量可接受的温室气体限值来避免气候变化的灾难性影响），然后为污染企业分配或拍卖等量的排放额度，企业将被允许自由使用这些额度，或在开放市场中对其进行交易。在“基线与信用”体系中，具体的绩效目标（也被称为相对常态估计的“名义基准”）会被赋予各污染企业，它们可以通过使实际排放低于指标来产生可交易的信用额度。

这些机制的支持者认为，交易体系比自上而下的监管政策（比方说设置碳排放税）更具效率和灵活性，因为它们会对公司内在的创新动力加以利用。与以同样方式影响所有行业的全面监管不同，他们认为减排市场是基于独创性和比较利益而建立的，能激励企业一同遵从守则。那些创新速度更快、效率更高的公司可以将他们获得的许可额度出售给创新能力较差的企业，从而使它们有更多时间迎头赶上，这样就能灵活而渐进地向符合内部市场动态的低碳经济转型。此外，交易将使企业家能够通过决定“其经营成本效益的削减程度，以及如何减少排放量以降低获得许可的成本”来“自由选择如何开展排污工作”。这样将确保排放量能在最具成本效益的情况下得到降低，且碳排放的明码标价能有组织地从市场内部产生，而不是从外部强加而来。支持者还设想，交易的方案能降低监管成本，因为市场一旦建立，就能根据其内部供求关系运行。而且，他们认为这些制度能减少“规制俘虏”现象产生的危险，即私人利益控制公共监督机构的过程，因为在交易计划中市场基本上自我控制的。

尽管排放交易计划有一定的优点，但它已经演变成了不稳定且具有潜在危险的机制，这些因素的影响实际上已超过了其大部分（如果不是全部）假定的优势。在一些情况下，它们只是将气候变化市场化，使之成为另一种投机和引起金融风险的机会。欧洲可能是碳市场的脆弱性表现最强烈的地方，世界上最大的碳排放交易市场就在这里。欧盟碳排放交易体系（ETS）于2005年启动，是欧盟广泛宣传的应对气候变化方针的奠基支柱之一。该体系覆盖了28个欧盟成员国，以及挪威、冰岛和列支敦士登，目前已经进入了第三阶段，该阶段将持续到2020年，届时欧盟将在联合国支持的现行议定书下达成其减排目标。欧盟碳排放交易体系覆盖了超过11000家工厂、发电站和其他类型的设施（共占欧洲总排放量的40%），在2012年1月，其覆盖范围又延伸到了民用航空领域。尽管该体系已被欧盟当局列为全球最佳实践，但自从其启动以来，也一直存在着严重缺陷。可交易配额的初始分配（第一阶段，2005—2007年）中目标宽松，给予了强势利益集团更多的配额和过高许可。关于化石燃料工业的激烈游说在欧盟总部所在地布鲁塞尔和欧洲各国首都经常发生，各个国家的实际配额正处于研究决定之中。根据“开放欧洲”智库的统计，欧洲各国政府“在2005年许可的二氧化碳排放量为18.29亿吨，实际排放量为17.85亿吨”。由于免费配额的确定是基于每个行业的历史排放量（这个过程被称为不受新法规约束的“祖父条款”），并会根据未来规划进行估量，该计划也给了欧洲污染最严重的企业意外的收获，难免会给执行计划较差的企业带来好处。德国是欧洲污染最严重的国家，也是最大的碳排放市场，其环境

部长指责该国四大能源公司——意昂集团（Eon）、莱茵集团（RWE）、大瀑布公司（Vattenfall）和巴登符腾堡能源集团（EnBW）从碳排放交易体系中牟取暴利，牺牲了消费者的利益，它们2006年的收益竟高达80亿欧元。“绿色和平”环境保护组织将欧盟碳排放交易体系称为“污染者的印钞许可证”，认为依靠对未来的排放量估测（这可以轻而易举地被行业夸大）会导致免费发放排放权，而企业之后会将权力转让以获得利益。

2006年，当宽松的目标和过高配额被证实，排放配额价格在几天内从每吨约30欧元（被认为是达到减排目标的最低标准）暴跌至每吨9欧元。之后，在2007年年中，排放许可的名义价值直线下跌，碳市场停滞不前。据安永会计师事务所（Ernst & Young）统计，欧盟碳排放交易体系造成了碳价格的波动，而不是鼓励了对可再生能源的可持续投资。与其所谓的目标相反，“该计划鼓励了短期的头寸交易，优化了回报，最大限度地降低财务风险”。

排放市场的可持续性也受到新政策的影响。例如，当欧盟出售额外许可来增加资助绿色能源计划的收入时，这难免会使已经膨胀的市场再次加大。而且，经过修订的能源效率法规也能引起减排，但这并没有纳入原本的限额，致使碳价格下降。由于价格本质上是不稳定的（因为它们会受到各种市场和非市场动态的影响），投资者预计，在2020年第三阶段结束之前，相对该年18亿吨许可的计划上限，其多余的许可将达到8.45亿吨。

在2009年提交给英国议会的备忘录中，剑桥大学电力政策研究小组的研究主任戴维·纽伯里（David Newbery）证实，欧盟碳排放交易体系“不能为长期低碳投资决定提供可预测和稳定的碳价格”。2004年，英国贸易和工业大臣帕特里夏·休伊特（Patricia Hewitt）已经写信给当时的欧盟委员会主席罗马诺·普罗迪（Romano Prodi），对欧盟成员国建立和管理碳排放交易体系的方式进行申诉。她警告道，“超出需要的分配实际上是在给企业一种自由资产”，而且“过度分配意味着交易极少或完全不发生，这种风险非常现实”，所以“贸易机制的可信度可能会受到损害，而欧盟及其成员国则需要寻找其他灵活度较低和成本更高的管理手段来履行其义务”。随着第二阶段（2008—2012年）新许可的出售，碳排放配额的价格再次在30欧元左右徘徊了一段时间，随后下滑到不到10欧元。

2012年，欧洲碳排放市场的价值下降了三分之一（从1480亿美元下降到1000亿美元左右），这增加了欧洲政府提供额外支持的压力。2012年12月，欧盟以

645万欧元的价格出售了558万份碳排放许可（作为第三阶段的一部分），对于企业来说，这个价格低到让它们无法对低碳经济进行严肃的投资。正如欧盟委员会最近所认识到的那样，碳排放权价格暴跌实际上可能会扭转减排趋势，转而导致对高排放技术的投资。此外，价格的持续走低也改变了市场投资者的看法，他们不再愿意以每吨10欧元的价格购买配额。最后，在2013年4月16日，当欧洲议会否决了未来几年减少碳信用额度的提议时，碳排放权价格在十分钟内下降了约50%，从5欧元下降到了2.63欧元。所有分析师都认为，在这种价位下，欧盟碳排放交易体系实际上阻碍了对替代能源的投资，使投资者对于“高污染”行业的支持更加有利可图。

然而，欧盟碳排放交易体系的灾难并不只体现在价格崩溃上。当2009年中期，市场成交量达到顶峰，有几亿吨配额的交易价下降到了每吨12欧元左右时，欧洲刑警组织（欧盟的执法机构）开始了一项调查，结果发现，“排放交易市场总量的90%是由欺诈活动造成的，这大大破坏了欧盟碳排放交易体系的发展”。超过一百人因“不明交易”罪名（一种逃避增值税的形式）而被捕，在欧洲十一个国家中，税收损失量约为50亿欧元。欧洲刑警组织报告强调了该体系的两个主要问题：一是碳排放市场的无形性（这使它们类似于投机金融市场，政府当局在其中对货币流通的控制和追踪能力有限），二是碳排放权交易商的登记程序（比较松懈，主要基于自我监督，从而增加了洗钱的风险）。

欧盟碳排放交易体系已被证明非常容易受到内部冲击、投机和有组织犯罪的影响。一系列“网络钓鱼攻击”，包括提示用户泄露其身份代码的电子邮件，导致了2010年整个欧洲国家注册机构的关闭。金融黑客也能顺利登入公司服务器，并在“现货”市场上出售配额，允许许可的即时交易以换取现金。现货市场比2008年增长了450%，总计达到了14亿吨，到2009年，现货量增长了75倍。2010年在德国发生了第一起“盗窃”事件，但政府当局却对此视而不见。之后在2011年，来自好几个国家总计3000多万欧元的配额被盗，这约占现货市场的20%，从而导致市场关闭了好几天。从那时起，政府才开始要求进行彻底调查。

欧盟碳排放交易体系的诸多缺陷令环保组织感到震惊，他们指出，欧盟在抗击气候变化方面的领导地位可能是审计计谋带来的结果，这种计谋被在现实中无意义的数据所掩盖：金融方面的运作并不能体现在实际减排目标的趋势中。正如世界自然基金会指出的那样，“信用额度太高，这破坏了欧洲的排放交易制度，会

允许欧盟在继续排放的同时仍然声称要达到减排目标。这可能意味着欧洲正在积极使气候问题恶化，而非改善它”。同样地，“开放欧洲”智库也将欧盟碳排放交易体系描述为“环境与经济的失败”。

其他以市场为基础的应用，特别是碳补偿机制，使这些问题更加复杂。这些机制进一步加大了对抗气候变化的风险和曲折程度。碳补偿允许企业和个人为他们在其他地方制造的温室气体做出赔偿抵消。抵消量以二氧化碳当量公制吨

（CO<sub>2</sub>e）进行量化，由官方机构认证，然后以“信用”的形式进行交易，就像在传统排放交易计划中一样。自2005年《京都议定书》签订以来，补偿行业就开始呈现指数级增长。《京都议定书》建立的国际框架“清洁发展机制”（CDM）是世界上地理范围和覆盖量最大的补偿方案，也是第二大碳市场。迄今为止，该机制已经发布了11.4亿美元信用额。2007年，其估价超过了330亿美元。然而，就像欧盟碳排放交易体系一样，清洁发展机制抵消额的价格在2012年同比下降了90%，达到了40美分左右，比“分析师所言中开发商获得信用额度的所需付费”大概低了10美分，“且远低于为减排设备投资的成本”。如今，只需点击数千个专用网站中的任意一个并通过信用卡付款，就能很容易地补偿在日常通勤、度假和其他任何形式经济活动中产生的排放量。清理我们的生态足迹从未如此简单而便宜。

为了防止碳补偿市场产生任何不良后果，认证过程绝对是最重要的。如果一项补偿声称在一定程度上减少了温室气体排放量，这也需要真实地反映出来。但是，测定“基准情景”的基线，也就是在没有补偿项目的情况下，测量什么是排放情景，这会受到许多方法性和概念性的挑战。如何区分一定会发生的减排和只有通过补偿计划才能实现的减排？这就要引入专家们所说的“额外性”原则：真正的补偿必须是对市场激励的反应，因此必须在正常发生事件的基础上产生更多抵补。人们可以通过两种方法来确定是否具有额外性：一种是对特定项目的方法，另一种是标准化方法。针对特定项目的方法是基于对补偿计划拟定产出的评估，然后在没有碳市场的情况下，从最可行和最可能的情况（针对各种检查程序）的产出中扣除。相对地，标准化方法只是根据预先确定的一套标准来评估补偿项目，这些标准通常会要求项目不能是法律授权的，必须包含特定的预先核准技术，并且排放率必须低于大多数的同类型项目。

碳补偿市场的另一个基本要求是“永久性”：想要达成真正的补偿，减排量必须是永久性的。事实上，如果直接将温室气体排放到大气中，其总量将会增长而

不是减少。永久性对补偿计划来说是非常重要的，对那些可能随着时间推移而恢复的排放隔离项目（通常占了大多数）来说，尤其重要。一个相关的例子是再造林（重建原有森林）或新造林（建立新森林）。虽然广泛的共识是种植树木有助于减排，但从中长期来看，森林衰退和火灾发生将导致更多的排放物重新进入大气，从而抵消了设想的补偿效益。同样，地质封存（如碳捕获和储存）也有泄漏的危险，特别是考虑到地下动力的可能性。补偿行业的审计师制定了各种基于市场的机制来衡量和预防逆转的风险，包括购买专门设计的保险、发展信贷储备“缓冲池”和发行未来需要重新认证或更换的临时信贷。然而，所有这些额外的保障都会增加成本、减少交易量。

由于标准、基准和量化方法，整个补偿行业依赖于所谓的指定经营实体（DOE）的第三方认证，即由专业审核员进行评估和测量，证明补偿项目的可靠性，以便它们投入碳排放和交易市场。就清洁发展机制而言，这些认证工作均由独立审计机构完成，它们会“对项目建议书进行批准，查证已执行的项目是否已经实现了计划中的温室气体减排”。

尽管指定经营实体的认证起着根本性的作用，但这些认证标准也相当通用（例如它们必须在财务报告体系上拥有足够的人力资源和经验）。作为指定经营实体的唯一限制是公司不能有任何待解决的“渎职、舞弊或其他不符合其作为指定经营实体功能的司法程序”。为了维护公正性，指定经营实体被要求“以可信、独立、非歧视和透明的方式”开展工作，在公司不同部门为不同客户服务的情况下，必须“明确界定与其他部门的关系，表明部门之间不存在任何利益冲突”，以及公司不涉及“任何可能影响其判断或危及其履行职能的独立性和完整性的商业、金融或其他进程”。当利益冲突发生时，指定经营实体需要对如何处理这些冲突做出说明。它们也需要在整个管理结构中提倡公正性“文化”，并在网站上宣传其政策。对指定经营实体的唯一制度要求是它们需要建立一个向最高管理层报告的内部“公正委员会”。

基本上，这些机构需要进行自我监督。政府当局认为，由于潜在的公正文化，审核员的工作不会受到不正当影响。还是要强调，气候变化治理的总体原则是自我调控。数据模型和量化程序主要由金融银行和审计公司开发，它们会给人营造一种中立性的假象。与信用评级或成本效益分析中的贴现率一样，对额外性、永久性和未来逆转风险的评估最终是一种主观评估。当主观性在金融资源增



长的领域占据主导时，就会发生利益冲突。指定经营实体就很容易在客户行贿下，提供扭曲的认证。由于大多数指定经营实体也是财务审计人员，他们可能会发现自己在为客户进行查证的同时，也在为其提供其他类型的咨询服务。部门内的相熟与共同网络也为指定经营实体及其同行创造了危险的“熟悉性”。最后，威胁总是有可能存在的，无论是公开还是暗中威胁。据专业服务公司德勤会计师事务所（与安永会计师事务所、毕马威会计师事务所、普华永道会计师事务所并称为全球四大审计公司，它们均参与了补偿市场）报告，由于“缺乏对核查人员和碳补偿提供者的监管和执法机构，在自发性的碳市场中存在很高的欺诈风险”。普华永道（PWC）气候服务专家也指出，“在项目启动和运行之前”，存在“项目倡导者试图筹集资金”的情况，这意味着他们会最终“在信用实际存在之前便进行出售”。此外，由于“没有全球碳信用额度登记”，企业“可以将相同的信用额度出售给几个不同的市场”。

由于缺乏可靠的外部执行制度来保证指定经营实体的公正性，以及他们评估中必然会具有的主观性，审计不善和行为不检的案例也就不足为奇了。2007年，世界自然基金会委托进行的一项研究对独立审计师的角色提出了质疑，研究发现，许多清洁发展机制项目质量较差，且并没有真正带来减排。2010年，一项新的研究发现，评估人员对发展中国家900多个补偿项目的评估工作并没有改进。在A（最好）到F（最差）的范围内，它们所获的最高等级（仅由一个评估者作出）只是D。该报告总结道，提供额外性的尝试可悲地失败了：“由于项目评估的缺陷，可能会颁发大量不具额外性的二氧化碳排放证书。这就可能导致全球排放量的增加，与该体系预期的减排目标背道而驰”。

仅在2010年，联合国气候变化秘书处就因为发现了不法行为的证据，而被迫中止四个指定经营实体。其中一家公司是气候审计领域的巨头，德国的TüV南德意志集团（Tüv Süd）。它所负责查证的3.95亿吨补偿额中，有21%并未遵循正当程序。并且，该公司“在正在对项目额外性进行考量的情况下，依然给予了一些项目批准的意见”。有趣的是，该公司是2010年世界自然基金会报告中唯一的D级审核员。根据碳点机构（Point Carbon）2012年进行的一项调查，清洁发展机制投资的前景“黯淡”。大多数的补偿项目都在中国和印度展开，而据分析人员报告，其中存在越来越多腐败和欺诈行为。此外，有许多投资者“计划减少或完全停止投资清洁发展机制项目”。清洁发展机制政策对话高级小组发表的报告也认识到，补偿市场是“危险的”。他们承认抵消额价格的下跌（仅在2012年下跌就达到了

70%)，项目也在之后几年进一步削减。公共和私人投资者都“正在对清洁发展机制市场失去信心”，认为减排目标“微不足道，不再对国际私人投资具有强大激励”，从而削弱了“全球碳市场的技术能力”。

## 结论：当数据成为危险的干扰

欧盟碳排放交易体系和补偿机制都表明，现实情况与支撑市场政策的数据大不相同。与之映照的是，信用评级机构、权力地位、信息不对称和利益冲突方面这种情况的例子在现实生活中普遍存在。强大的公司利益影响游戏规则的能力是社会其他部门无法比拟的，无论是公共利益团体还是非政府组织。可再生能源公司的崛起也还没有改变化石燃料行业的市场主导地位。据“透明国际”组织透露，2009年美国石油和天然气公司的游说投资超过了清洁能源行业八倍。在欧盟，企业集团的政策立场重要性大大超过了环保集团。就像金融部门一样，新的碳市场也容易出现各种各样的差错，包括欺诈和犯罪活动。一些分析师估计，到2020年，减缓全球气候变化的总投资将达到7000亿美元的惊人数值，预计每年的公共投资至少为2500亿美元。在这样一个巨大的金融市场上，也会包含非常高的违法风险，特别是由于“许多气候问题的复杂性、不确定性和新颖性”，包括当“测量碳补偿环境保全的工具相对未经检验”时，“对森林的界定或者对额外性的确定”都有很大的复杂性、不确定性和新颖性。

带动气候变化否认主义的各种角色已经由各种动机激发。可以说，化石燃料工业的主要目标是通过指导政策议程来抓住权力。通过向政策制定者提供及时的反证，石油和煤炭联合企业在为大部分竞选活动提供资金的同时，实际上已经挟持了美国的政治制度，同时将其触手伸向世界其他地区。如今，从中国到俄罗斯，到巴西，没有哪个国家可以免于化石燃料产业带来的巨大压力。对于这个强大的企业联合体来说，任何对环境法规的延迟（或阻碍）都意味着长期的特权和优势。对于持怀疑态度的科学家来说，在资金和声望方面他们已经获得了回报。当关于气候问题的争议随时间不断增多，这些非正统的“专家”大部分都享有了前所未有的媒体曝光度。他们的论文、书籍和电影[例如电影纪录片《冷静一点》（Cool It），包含了隆伯格的大量观点]开始受到广泛的受众群体的欢迎。当你得到世界上最强大和最富有的公司的支持时，反对主流变成了一件非常值得做的事情。气候变化的倡导者在集体科学研究的海洋中畅游，因为他们的个人贡献只是一个巨大拼图中的一小块，这使得他们在很大程度上并不为人所知，而反对派人士却能在一个逆向宣传的小池塘中享受舒适的位置。在一个正常的世界里，怀疑主义应该被放逐到文化古玩店，就像曾经的神创论和“否认大屠杀”一样。但

是，当期刊、报纸和电视都开始陷入争议陷阱时，池塘升格成了海洋。两派人士拥有着相同的曝光时长和接近公众的机会。这给他们带来了巨大的利益，曝光度可以转化为薪金支票、演讲费和版税。并且，在市场化的学术界，名声能直接让你拥有一份顶尖大学的工作。

邓拉普和麦克里特坚持认为，不要在意这些团体和个人背后的各种动机，“是对政府监管力量的共同反对，将他们团结在了一起”。虽然反对派科学家的主张总是会随着时间演变，但“不需要监管”的主旨始终保持不变：“对自由市场的坚定保证和对政府监管的鄙夷，反映出保守的政治意识形态，这种意识形态在否认气候变化的团体里几乎是普遍的”。

功利主义推理一直是他们最有力的武器。尽管他们的科学主张（例如气候变化是一种骗局，它不是由人类造成的，对地球来说可能并不是那么糟糕）一次又一次被证明是错误的，他们的计量经济模型仍然成为气候变化经济学的主流。成本效益分析和基于市场的方法手段现在是环境政策制定的奠基支柱；贴现是公认的边际效用货币化方法；碳交易被广泛认为是对气候变化最具成本效益的回应。环境会计已经成为价值数十亿美元的新兴产业，补偿计划在越来越多的工业化国家迅速兴起。数以千计的公司（以及数百万善意的人）每天都会购买和出售碳信用额度。谁是这些新兴金融市场的主要受益者？主要是化石燃料公司和银行，这些为全球气候和金融危机负责的机构。污染工业通过排污权交易（主要在“主张进步的”欧洲）获得了意外收益，而投资和商业银行负责了大多数（如果并非全部）现有交易设施的设计。与此同时，当全球的公共机构处于前所未有的财务困境时，碳信用额暴跌至离谱的价格，给政府带来了巨大的税收损失，并产生了对额外公共资源的需求（达到一种“碳救助”的形式）。

有人认为，这至少会有助于减少二氧化碳排放量，但事实真的是这样吗？这其中仍然包括了许多疑问，因为会计体系相对粗略，最终容易引起争议。例如，国际能源署（IEA）等国际机构公布的数据是由污染行业（如石油、煤炭和天然气公司，水泥公司，建筑行业，制造业等）收集的，作为其碳排放报告要求的一部分，这会引起大家对最终数据准确性和有效性的怀疑。无论如何，根据全球大气研究排放数据库，过去十年来全球大气变暖的主要原因正是二氧化碳持续增长的排放。在2012年（根据撰写本文时提供的最新报告），全球二氧化碳排放增长了3%，“达到了340亿吨的历史新高”。据估计，目前大气中“由于人类活动累计排

放的”二氧化碳总量达到了4200亿吨。

成本效益分析的表面中立性隐藏了一个重要的潜在事实。尽管经济推理倾向于对所有成本做平等分配，但在现实生活中，这并不一定是正确的。成本花费在社会上并不是均匀分布的。一般来说，人们会认为按比例分配成本的决策是公平的。但在某些情况下，将成本只集中在某些类别的费用上可能更合适。在选择负担成本的两种方法时，如何对成本进行分配可能比整体花费量更为重要。例如，社会可以合理地选择一个比较昂贵的治理方案，其花费能进行平均分配（就比例性和功能性而言），而非选择一个便宜但会被认为不公平的方案。如果某些行业从传统意义上和竞争优势中受益更多，那么要求它们承担向低碳（或理想情况下的无碳）经济转化的大部分花费是公平合理的。而相比之下，成本效益分析则统一分摊社会成本，最终更强调采用“最便宜”的替代方案。这种推理模式总会为商业提供更多回报。

就像大多数交易计划一样，欧盟碳排放交易体系可能是目前最具成本效益的选择。但这种成本效益是对谁而言的呢？当然是化石燃料工业：正如我们所看到的，该体系导致了公共资源向私营部门的再分配。但是，对于公民来说，该体系是否同样能表现出成本效益呢？可能并不会。因为大多数的能源公司只会向其客户的能源账单收取更高费用，以购买排放许可。虽然欧盟碳排放交易体系在商业层面产生了相当大优势，却对日常预算产生了负面影响。

在环境治理理念的斗争中，怀疑论者最终获得了胜利。他们的成本效益分析把所有事情都转化为了数据，数据又转变为价格，能对我们的社会产生长期影响。这对国际社会来说是一种危险的干扰，它会导致世界陷入市场霸权和生态崩溃的恶性循环。尽管自20世纪60年代以来，就存在人类行为造成气候变化的证据（以及20世纪80年代后期形成的科学共识），世界还是一直在等待。等待经济学的手段，会计师的测量，以及审计师的查证。但是，我们似乎遗忘了，这种情况并不遵循经济模型，而且，更重要的是，并不能在审计师的检查表单内进行批准。

本章集中讨论了数据如何为缓解气候变化的市场化方法铺平道路，我们在下一章中将深入探讨环境治理的另一个关键部分，即自然资本和生态系统服务的估值。这两个方面确实是密切相关的。我们可能会说，如果气候变化的市场化现在已经通过排放权交易和补偿成为现实，那么为保护自然而引入类似的“市场”可能

会最终导致自然世界的金融化。与碳市场的一样，数据模型会让我们觉得，可以为任何东西定价。但是在试图为自然定价时，即使我们真正的意图是要保护自然，我们这一代人也正踏上了由投机市场、投资企业和私人审计公司组成的危险地带。

---

## 第四章

# 衡量不可测：数据把自然金融化

---

“衡量无法衡量的东西是荒谬的，只不过是一种从先入的概念进而导出预定结论的手段而已。不过，逻辑上的荒谬还不是这种做法的最大缺陷，更严重的和对文明起破坏作用的是妄认为每件东西都有价格，也就是妄认为金钱是一切价值中价值最高的。”

——E. F. 舒马赫

《小的是美好的：一本把人当回事的经济学著作》，1973年

“事物唯一的共同衡量标准是金钱。”

——杰里米·边沁

功利主义哲学家、经济学家

经济推理和市场手段的传播并不局限于争议性的减缓气候变化领域。它已经成为我们与整个大自然相处的主导趋势。过去几年里，非市场现象货币化的方法在两个领域里出现了意外而空前的增长：自然资本和生态系统服务。这两个核算领域的区别相当虚假。对它的一种描述是，自然资本核算旨在衡量常规经济过程（以GDP衡量）中环境“资本”的消耗，而生态系统核算试图衡量一般自然服务的经济价值。作为一个经济研究领域，自然资本/生态系统核算已经出现在评估人类经济体系与自然之间复杂相互作用的需要之上，后者经常为人类提供必要服务，却不在产权和市场动力学的架构内。



正如我们在第一章中看到的那样，国内生产总值所依据的国民收入账户忽视了一些非市场商品的经济价值，包括无报酬的工作、礼物经济、各种形式的“产消合一”（即人们生产自己的消费品）以及自然资源。批评者指出，经济表现的官方

衡量指标将自然资源视为可支配收入，但它们实际上应该被视为不可再生的资本，如果经济增长的过程继续下去，自然资源必须予以补充。他们的论点是，如果大自然需要被正确地评估，并纳入经济表现的计算中，我们必须认识到其贡献对经济福利起到的决定作用。这种认知会导致完全不同的宏观经济政策，环境保护就会在其中占据中心位置。考虑到这种情况，自20世纪80年代以来，经济绩效指标领域的先驱们一直在研究制定“真正进步”的衡量方式和“绿色”会计制度。他们已经采取了各种方法，来将经济过程中消耗的自然不可再生资源的价值纳入考虑范围，以便在国内生产总值中扣除这些数值，就像将资本消耗从国内产品的“净”估值中减去一样。1992年，美国经济分析局开始研究一系列卫星账户<sup>[1]</sup>，以捕捉地下矿产资源的经济影响。然而两年后，美国政府停止了这一举措，因为他们的一些发现遭到了采掘业的攻击，他们还被要求进行外部评估。一个由威廉·诺德豪斯担任主席的高级小组对此作出总结，认为衡量自然资源和环境是“一项重要的目标”，该小组由国民收入核算领域一些最著名的专家组成，其中还包括了罗伯特·艾斯纳（Robert Eisner），他曾在20世纪80年代提出了“总收入”账户制。在该小组1999年发表的题为《自然的数据》的最终报告中，他们建议采取分阶段的方法，主要侧重于建立森林账户，然后转向农业资产、渔业和水资源。尽管如此，他们也认识到了开发这些新账户所涉及的方法性挑战：“该过程将需要解决重要的概念性问题，制定合适的物理测量法，并对相关流量和存量进行估值”。

总的来说，这些核算制度旨在实现双重目标：一方面是考虑到自然资源和其他环境服务的不可持续消耗带来的收入，对国民经济核算的国际体系进行改革；另一方面是要提供关于自然生态系统经济价值的信息，以强调大自然对人类福祉和经济增长的重要性。正如经济学家西蒙·希克斯（Simon Hicks）在20世纪40年代所说的，任何对收入的衡量，测量的都应该是个人（或社会）在不损害未来消费能力的情况下创造财富的能力。但是，如果目前的收入衡量标准主要基于不可再生的生态系统资源的不断消耗，那么从长远来看，这种模式是无法持续的。罗伯特·雷佩托（Robert Repetto）和他的团队在1989年进行了关于如何将自然资源纳入国民收入账户的初步研究，用他们的话来说，“一个国家可能会耗尽矿产资源、砍伐森林、侵蚀土壤、污染地下蓄水层、过度狩猎野生动物、破坏渔业资源，但是，虽然这些资源消失了，对收入的估量却不会受到影响”。根据2013年一项衡量世界真正进步的研究，全球GDP增长与1978年以前人类福利的总体改善息息相关。在那之后，经济增长（总体水平）导致了人类、社会和自然均衡的恶化。



自然资本能体现在人类福祉的很多种方面，而官方GDP统计数据却经常忽略这一点。自然资本产出的物品能在市场上流通，就像农业部门的大多数产品一样。它还产出有助于经济增长和人类福祉的生态服务和便利设施，例如供水、土壤施肥、授粉等，但这些都不是通过市场引导的（例如自然为我们提供水源是完全免费的）。同时，也有经济过程会对人类福利和自然资本都产生负面影响。废物的产生、污染和有害物质富集都是经济增长带来的后果，它们对环境不利，对人体有害。由于人类福利不仅仅只是对经济产品和服务的消费，自然界所发挥的基本作用确实不能忽视。但是，大自然可以被定价吗？定价的过程中会出现哪些固有的方法和概念性问题？会不会有其他的政治议程隐藏在这样一种表面良性，为自然“估价”的尝试中呢？本章着眼于数据如何影响全球关于生物多样性保护的辩论。我们会讨论到将自然转化为数据的关键方法的历史性演变，并剖析现有方法的优缺点。由于本书主要关注数据在治理中的社会作用，本章认为自然资本核算和生态系统服务领域的这些新趋势有可能会失控。鉴于投资银行、私人审计师和各种各样的咨询公司会对这些新的“市场”产生兴趣，因此存在这样一种严重风险，那就是一开始出于好意的工作（如为自然“定价”以更好地保护它）实际上可能会导致世界自然奇观的金融化。

## 衡量自然的价值：全球治理的数据化演变

2007年，八国集团首脑会议在德国的海利根达姆举行。其中，他们签署了环境部长提交的一项提案，即“启动分析生物多样性的全球经济利益、生物多样性损失的代价以及不采取保护措施与有效保护的成本之间的对比”。他们希望通过无可争辩的数据，对环境治理目标进行清晰、合理的成本效益分析，以反映出准确的价格。不是概念，不是通用的目标，而是以市场价值形式出现的粗略统计数据，将会更容易融入经济计划，并引起社会对经济增长的广泛关注。这一举措促成了欧盟委员会和德国政府就“生态系统和生物多样性经济学”（通常被简称为TEEB）的协调研究，对货币估值技术进行了整理，以便定义一个通用的政策应用框架。项目报告在介绍中承认了研究任务的复杂性和重要性：

我们仍然在努力寻找“自然的价值”。对于我们来说，自然每天都能创造出无数价值，但它们大多绕过市场，躲开定价，不被估量。我们渐渐发现，这种估值不足是造成生态系统退化和生物多样性丧失的根本原因。

随着对自然资本估值兴趣的增加，有更多机构加入进来。2010年，世界银行发起制定了一项合作机制“财富核算与生态系统服务价值评估”（WAVES），这是一项全球性伙伴关系，旨在就衡量自然资源经济价值的一般方法达成共识。该机制的主要目标是“确保自然资源纳入用于衡量和规划经济增长的国民经济核算的发展规划，以促进可持续发展”。它特别着重于建立公共和私营部门之间的协同关系，将对自然资本的估量纳入企业领域的战略选择。世界银行对自然资本的核算建立在其长达十年对“总财富值”的新测算上，这些新方法受到了“古典经济学家观点”的启发，“他们将土地、劳动力和生产性资本视为生产的主要因素”。这些研究中使用的关键指标之一就是所谓的调整后净储蓄，更被广泛称为真实储蓄，它通过考察国家提供的未来资源，来测定可持续发展的能力。联合国提出的包容性财富指数等财富总指标能够区分生产资本（包括机械、设备、基础设施和城市用地）、自然资本（包括土地资源、森林和地下资产）以及无形资本（包括人力资本、机构质量与治理等一系列资产）。在题为《国家财富在哪里？衡量21世纪的资本》的报告中，世界银行宣称，“在较贫穷的国家，自然资本比生产资本更重要”，因此，他们认为对自然资源的细心管理应该成为发展战略的一个基本组成部分，“特别是，这些国家中最贫穷家庭往往最依赖自然资源”。然而，必须强调的

是，这些类型的指标采用的是“弱”可持续性方法，这种方法认为，自然资源的消耗可以通过其他领域的等效投资抵消。因此，为建造幼儿园而拆除一个公园会产生“完美的平衡”，这对可持续发展没有负面影响，或者说，在经济方面，这么做对于后代的净机会成本等于零。在这种核算模式中，我们从自然中拿走的东西会由我们在人类发展中的投资而重新平衡。

然而，直到2012年，联合国统计委员会才通过了第一个自然资本估值的国际标准：环境经济核算体系（SEEA）。SEEA的起源可以追溯到20世纪80年代末和90年代初国民账户体系的修订，当时世界政治和经济舞台上的历史性变化（从冷战结束到全球化市场的增长）使联合国对会计准则进行了重要修订，并于1993年发布了新准则。当时，联合国审查人员拒绝了彻底修改国内生产总值的要求，因为他们不愿在国民账户中增加更多估算数据。他们批评环保倡导者试图将“规范性措施”强加于经济表现的计算中。同时，他们也坚持认为，由于自然资源并不是从大自然“购买”的，无论其估价如何，都难免是武断的，且具有争议性。因此，不做估值反而比扭曲地估值更好。他们将自然资源损耗与人造资产贬值进行了区分：和前者不同，后者指的商品消费，其生产已经在体系中做出充分考虑。联合国统计人员并不承认自然是生产的一个要素（符合主流的新古典主义经济思想），因此传统的国民账户不适合进行环境核算。出于同样的原因，2008年的下一轮改革放弃了将自然资本损耗纳入国民账户考虑的可能性。这次修订过程再次肯定了这样一个原则，即将一项活动视为生产活动的“必要条件”是“它必须在一些机构的授权、控制和负责的基础上进行，而这些机构对其生产的产品拥有所有权”。例如，由于没有专有机构管理公海鱼类的自然生长过程，鱼也不属于任何组织或公司，所以不受国际配额限制的公海鱼类生长不算作生产。相比之下，养鱼场鱼类的增长则视为一个生产过程，会加入国内生产总值的计算中。另外，“野生、未开垦的森林、野果或浆果的自然生长”也被排除在生产之外，而“作为作物、木材或其他用途的树木”，其计算方式与一年生作物的生长相同。同样，“对野生森林中树木的蓄意砍伐”以及“野果、浆果或木材的采集”都被视为生产。按照相同的逻辑，“降雨和水沿天然河道的流动”不是生产过程，而“在水库或水坝中蓄水，以及将水源从某地输送或运送至另一个地方”都会对国民收入构成正增长。

当在官方GDP统计数据中减去自然资本消耗的过程遇到来自联合国统计人员的阻力时，更简便的方法是将其纳入卫星账户系统，同时进行一系列计算，其中包括对非正式部门、旅游业（在GDP账户中只进行少量统计）和无偿工作指标的

衡量。正是在这个领域，SEEA才被开发为一种互补的账户体系。它为评估的标准化提供了一个总体框架，用以处理经济和环境之间物质和能源的直接物理流动、环境资产的存量，以及与环境有关的经济交易。如图4-1所示，SEEA将自然资本和生态系统视为生产要素，并衡量其对具体经济活动的贡献。环境经济核算体系以其语言、概念框架和方法论，通过将GDP核算延伸至评估自然对经济增长的贡献，反映其逻辑性。

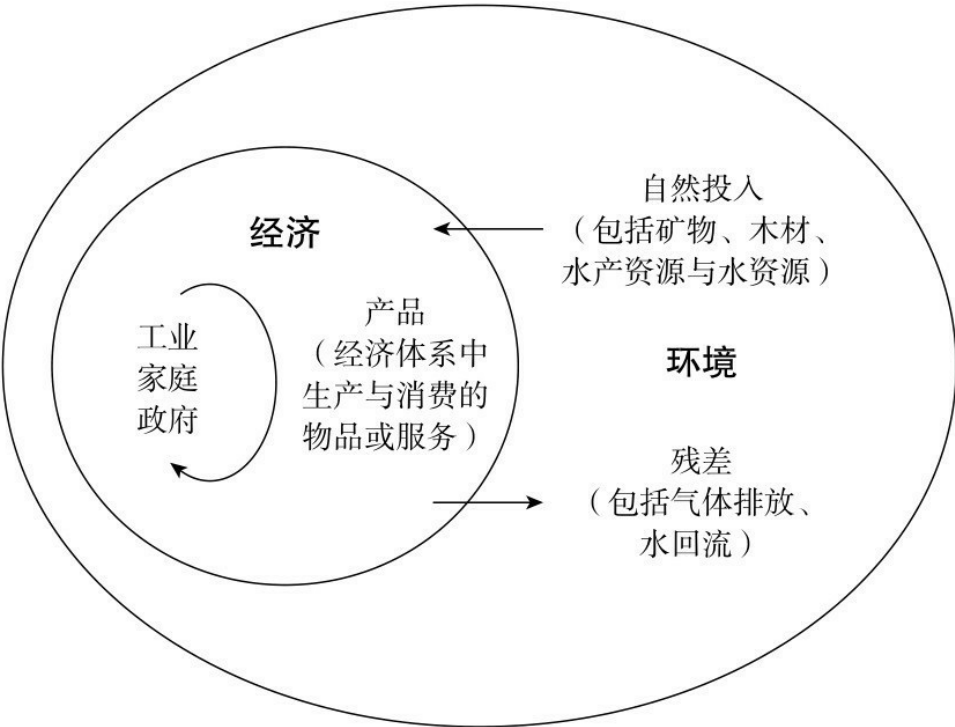


图4-1 自然对经济的贡献概念化

资料来源：联合国统计委员会等。

由于SEEA关注的是，直接使用环境资产作为经济的自然投入所带来的物质利益，所以它并不是一个综合的生态系统服务评估框架。SEEA仅为国民收入统计资产范围内的可再生和不可再生自然资源以及土地的估值提供“指导”，这些估值也会反映在GDP上。它的目的不是衡量自然资产或是超出传统国民账户体系的相关存量和流量。在撰写本文时，联合国仍然在对生态系统评估方法的制定进行讨论。他们计划发表一份关于“实验性生态系统账户”的报告，该报告将涉及对全人类物质和非物质（以及直接和间接）利益的衡量。然而他们澄清道，这种衡量“并非一种统计标准，但它会对使用系统方法测量生态系统的现状进行连贯一致

的概述”。

在SEEA中，估值范围仅限于经济主体。在有明确生产和财产边界的GDP账户基础上，SEEA仅衡量经济所有者获得的收益，经济所有者指“有权通过其所承担的相关风险享有经济活动过程中资产使用带来的利益的机构单位”。就环境资产而言，这种利益以“销售自然资源和培育生物资源的营业盈余”形式记录，租金以“许可使用或开采自然资源”来赚取，或当环境资产被出售（例如出售土地）时，以“净收入”（即不包括交易费用）的形式表现。当这种“交易”不存在市场价格时，价值评估必须基于概念假设和数值模型，SEEA为此提出了一系列建议。一种方法是，只要有可能，就应该通过直接观察在实际市场上交易的资产价格，对未交易的类似资产进行定价。例如，土地、木材和矿物的销售信息可以用于估算尚未销售的类似商品的价值。另一种方法是着眼于重置成本，特别是对于不能或一般不进行交易的自然资产。根据这种方法，自然资产的价值相当于等价新资产的当前收购价格，扣除其折旧。就像二手车的价格相当于新车的成本减去其使用周期和里程中的贬值一样，“可以运用这一方法估算属于固定资产的培育生物资源的存量价值，例如果园的价值”。

当这两种方法都不能使用时（比方说没有相关的市场交易或收购价格），SEEA建议采用一种从数学上看来更复杂（但基本上会规避）的方法，它将自然流量的价格作为未来收益折现值的基础，正如我们在第三章中已经提到的传统成本效益分析一样。未来回报的折现价值通常被称为净现值（NPV），“使用资产将来的预计开采速度和它的预计价格，生成一个预计回报的时间序列”。按照传统的经济思想，即假定现在获得的回报高于将来获得的回报，预期回报流会被折现，以反映“买方准备为当期资产支付的价值”。此外，为了确保估价符合一般市场价格概念，联合国统计人员建议自然资本的审计员使用与生产资产假定回报率相等的市场折现率。对于净现值方法来说，实际上，没有预期回报的资产在经济上就没有价值。然而，由于预期回报在定义上仅仅是由估计和推测得出的，而非实际观察得出，从而使整个论证由统计核算（常常被描述为现有数据的中立性集合）转变为财务分析（这属于价值判断和风险评估领域）。但SEEA是在2008年金融危机之后编写的，作者已经意识到，“金融市场可能不够发达，不能提供这些特定回报率的可靠估值”。因此，在看上去更像是基于举例证据而非可靠经济分析的政策建议上，他们建议使用真实回报率，比方说“政府债券利率”。他们认为，拥有自然资源所产生的投资回报应该大致相当于主权债券的回报。显然，该报告一定

是在2011年发达经济体经历主权债券降级之前编写的。

除了这些模型的表面中立外，联合国统计委员会也认识到，所有这些方法都是有问题的，并且基本上是主观的。它也承认，该框架有意地避免了用价值计量“今世后代可能获得的所有收益”，所以并不采用“可能被视为环境资产的社会估价的东西”。另外，它关注的是特定的财产边界和对经济的直接投入，而忽略了自然资本对社会更广泛的影响。联合国统计人员认为，更好的方法是在环境资产评估中应用“社会折现率”，而非市场折现率，因为前者反映了“对整个社会广泛的长期价值”，而后者的计算“单纯依赖它们对当下利用者的价值”。因此，他们得出结论，支持使用社会折现率的一个主要观点是，它们“相对来说更重视未来世代的收入”，而使用市场折现率得出的净现值估值“不重视未来世代”。

## 生态系统的价格评估

1997年，以马里兰大学的生态经济学家罗伯特·科斯坦萨（Robert Costanza）为首的一个研究小组首次尝试对生态系统服务的总价值进行了估计，这个生态系统服务即自然界在世界范围内为经济活动无偿提供资源的全过程，包括供水、授粉等。在《自然》杂志发表的一篇论文中，他们收集了当时为止所有可用的研究成果，估计整个生物圈的整体价值约为每年（平均）33万亿美元。当时，全球GDP约为每年18万亿美元，而研究得出的生态系统服务总价值几乎达到了它的两倍。研究人员因此受到了极大的激励。当然，他们知道很多人会认为，“对生态系统估价既不可能，也不明智”，而且“我们不可能对诸如生命、环境美学或长期生态效益等‘无形的事物’赋予价值”。但他们的结论是，由于缺乏明确和透明的评估，生态系统将继续在政策决策过程中被赋予“很少的权重”，而最终危及“人类在生物圈中的可持续性发展”。2002年，在《科学》杂志上发表的一项后续研究中，他们得出的结论是，全球自然保护的净收益（以净现值衡量）远远超过成本，二者之比达到100比1。投入自然生态系统保护中的每一美元，都会产出100倍的经济效益。从投资角度来看，从没有金融泡沫或投机衍生品可以产生这么多钱。正如一些评论家所言，“就算是在9月11日之前令人陶醉的网络热潮时期，这样的数据也足以使每一个当日交易者感到眩晕”。但这些数据的背后是什么？

正如SEEA框架的演变所证明的那样，自然资本的估值方法充满了谬误和风险。当分析领域扩大到包括无穷无尽的各种生态系统服务时，这些问题就更加复杂化了，在这种情况下，就不仅仅要衡量某些部门对经济增长的直接贡献（如SEEA所做的那样），还要（且更重要）衡量它们对人类福祉的整体影响。科斯坦萨和他的同事们在关键研究中承认了“生成这种评估固有的概念性和经验问题”。他们的方法难免过于简单了。他们以地球上类似气候条件的地区划分生物群落，然后计算16种生物群落中17种资源设施的每公顷平均值，再将该值乘以地表总面积。当结果呈现给大家时，很多数据无法找到。例如，在他们的论文中找不到对沙漠、苔原、冰岩和农田等区域重要生物群落的估价研究。他们还指出，做估值会不可避免地导致“静态快照”的出现，而自然实际上是一个非常复杂、相互依赖的非线性系统。此外，货币化框架假定在“生态系统响应函数中没有明显的阈值、不连续性或不可逆性”，但事实“几乎完全不是这样”。

还有一些关键性的规模问题限制了生态系统服务货币化研究的政策和社会影响。以33万亿美元这个数据为例，对普通公民或决策者来说，这究竟是高还是低？即使是这样一个简单的问题，我们也很难回答。尽管有些人可能会难以理解如此巨额的金钱（就像他们难以捉摸全球经济的规模一样），但仔细考虑，这个数据实际上可能非常小。如果自然的价值相当于两倍的全球GDP，那么一些人可能会得出结论：在最坏的情况下，如果生态系统完全崩溃，那么发展三倍的世界经济就足以取代所有自然资本。如果所有的工人都能制造出三倍数量的产品，所有的公司都愿意承诺三倍的生产目标，那么人类就可以没有自然。我们可以成功将降雨、授粉、水资源等取代。这样就形成了什么都不缺，而且非常便宜的自然救助。这可能吗？我们中的许多人会在这样的想法面前退缩，因为我们会认为这些推理只是某种经济上的妄想，或者彻底的政治疯狂行为。两倍于全球经济？即使认为生态系统服务是全球经济的三倍、四倍、五十倍，仍然是它的无限大价值的愚蠢低估。就是这样。然而，无限大的计算永远不会引发政策改革。它只会成为一种通用的说法，而不会在实践中起什么作用。通常情况下，当数据比我们的日常经验所知大得多时，我们无法掌握它们的实际大小。当数据太大，无法理解时，就不太可能改变什么政策。但当它们相对较小时，就存在相反的风险。

由于这种不可避免的估值困境，许多评论者认为，将生态系统货币化的目标（或者我们常说的是，为无限大的价值制定价格）是不可能实现的，也是不明智的。在对生命、环境美学或长期生态效益等“无形资产”的经济价值功利计算中，有显而易见的可疑性。从研究小组在《自然》杂志发表论文以来，许多生态经济学家对生态系统服务评估方法的各种技术方面都提出了质疑，特别是对分散的调查和重置成本估算的依赖，而这种依赖会进一步受到重复计算风险的影响。

生态系统的货币化有多种方法。其中一些通过跟踪相关市场的价格来间接评估环境“货物”的价值。例如，用住房价格的特征方法来估计当地环境质量的价值。以两幢公寓为例，假设它们的大小和质量大致相同，但一幢位于公园附近（因此更昂贵），而另一幢不在公园附近。那么公园的估计价格是多少呢？要用基于两幢公寓之间价格差异的算法来回答这个问题。同样地，游客的旅游成本已被用作休闲度假场所价值的代表，而那些试图保护自己免受环境风险所花的费用（例如喝瓶装水、医疗开销等）也会被用于货币化生态保护的价值。然而，大多数的方法显得更“直接”，它们遵循所谓“支付意愿”（WTP）原则的概念框架，通过调查研究，试图基于潜在买家获取某些商品的意愿来估计价格。许多真正的环



保主义者并不赞成“支付意愿”原则，它是在市场研究中开发的，并最终被大公司普及，作为在新产品开发过程中估量决定价格的主流方法。公司通常采用“支付意愿”原则来制定最优的定价策略（从而打败竞争对手，增大市场份额），预测市场对价格变化的反应，对需求函数建模，并衡量品牌对特定产品的附加价值，与无品牌基准产品作对比。其根本原则是，当我们买东西的时候，我们的评估是建立在一系列“客观”和“主观”的参数上的。客观参数包括功能、新鲜感、实用性等。主观参数包括合意性、社会地位，毕竟并不是我们购买的所有东西都具有使用价值。此外，我们的购买行为并不是在封闭状态下进行的，而是以惯例和期望为特征的复杂社会动态的一部分。我们的支付意愿建立在所有这些参数上，当市场价格等于（或接近于）我们的标准时，我们可能会进行购买。这就是为什么公司都要尽可能多地了解消费者的偏好。他们会收集人们购物习惯的数据，他们想知道我们在上网时搜索的东西，以及我们喜爱的物品。如果他们能在我们的脑门上看到我们的支付意愿，就能设定相应的价格，使销售额猛增。然而，在他们的营销策略中，他们并不是简单地、被动地猜测我们愿意为特定商品支付多少钱。他们正在积极影响我们的心理估值过程，主要的方法是广告和品牌推广。他们希望我们的支付意愿能够反映他们的期望和利润率。在这种持续的相互交流中，衡量支付意愿既是一项启发式任务，也是一场心理战。

对支付意愿的测量方式大致可以分为两大类：基于显示偏好和基于陈述偏好。第一类方法中，我们发现在像维克瑞（Vickrey）拍卖<sup>[2]</sup>这样的实验中，参与者在不知道其他出价人估价的情况下提交密封投标，与在线委托代理拍卖很相似，尤其是eBay推广的拍卖类型。在这些实验中，成功的投标人会被要求以投标价格购买商品，尽管实际价格会被设于第二高价（第二价格拍卖）。还有一项类似的实验是让参与者同时提交报价，其中销售价格从零到一般商定的最高价格之间随机抽取。出价大于销售价格的投标人会收到一个商品，但只支付相当于销售价格的金額。这两种拍卖机制都旨在鼓励参与者展示他们真实的估值，这是经济学家普遍定义的激励相容性原则：由于他们对拍卖商品的购得不是基于最高投标价格，参与者就不会对拍卖过程进行投机冒险（出价过高）。

在第二类方法中，我们发现了基于调查的测量方式，参与者被要求对某些项目（如偏好结构）的支付意愿进行排名，或直接确定最小—最大的价格幅度。在某些情况下，有限的资源和时间限制会使焦点小组更适合运用自下而上的研究方法，例如实验和调查。在这方面，专家判断成为“评估顾客支付意愿的启发式方

法，并提供对不同价格水平的通用需求估计”。事实上，人们普遍认为，衡量支付意愿的调查手段存在根本性的问题和不确定性。受访者并没有受到激励来展示他们真正的喜好，还会有一些其他的考虑在其中发挥作用。例如，有些人可能会因为“声誉效应”而倾向于夸大价格，会以此不让自己显得“小气”。相反，他们也可能试图人为地降低价格，因为很多人认为他们的角色是“认真的买家，应该帮助起到调低价格的作用”。买家几乎没有估计产品价格的能力，特别是当某种购买并非高频率时，一旦受访者知道了实际的市场价格，他们填写的支付意愿就会突然发生变化。对个人的调查往往会导致矛盾，甚至相反的结果。此外，市场研究人员也意识到，假设定价与实际购买之间存在着不可调和的分歧：即使受访者透露了他们的支付意愿，这也无法自动转化为实际的购买行为。在很多情况下，决定产品的具体价格是否可以接受会比直接指定价格来得容易。那些显然不重要的方法工具，例如封闭式问题（受访者在调查者列出的价格中进行选择）或开放式问题（受访者必须提出自己的估计值），会产生截然不同的结果。不出所料的是，封闭式二元选项问题（也就是只有两个选项的二选一问题）更容易产生可靠的结果。在间接调查中，同样也会出现这种情况。受访者会被展示不同价格的商品外形，并被要求表明他们是否会以该价格购买商品。这种调查在受访者中呈现出了非常不一致的结果。

估算支付意愿的各种困难，在传统营销领域都很常见。特别是当市场不存在的时候，情况会变得更加复杂，就比如生态服务系统。在这种情况下，估算过程必须通过所谓的条件价值评估法进行模拟：并非通过设定价格来估计买家的支付意愿，而是用支付意愿来估量价格。因此，如果没有实际价格来检验调查结果，那么人们愿意支付的价格就会成为自然环境事实上的价格。首次发布该方法的参考文献出现在20世纪40年代后期，研究人员在其中提到了基于支付意愿的调查可能性，以测量预防水土流失所能获得的金钱“回报”。然而直到20世纪60年代，条件价值评估法才开始被更加系统地用于各种主题的研究中，从自然保护区的价值评估，到狩猎水禽的权利研究，以及捕鸭许可证的价值研究等。条件价值评估法也开始在非环境问题的研究中受到应用，例如降低心脏病发作死亡风险的政策、预防呼吸系统疾病和改善杂货店的价格信息。

20世纪80年代，条件价值评估法也被应用于政策改革，特别是在自然灾害评估领域，引发了学术界和商界的激烈争论。一个有趣的案例是1980年《综合环境反应、补偿和责任法》的制定，它也被简称为CERCLA，更常见的名称是《超级

基金法案》。该法案规定，当作为自然资源托管人的公司对自然资源（包括湖泊、溪流、森林、海湾、沼泽、陆块等）造成破坏时，美国政府机构有权起诉它们。法案中还规定，条件价值评估法不仅可以用来估量对使用者造成的直接损害，还可以用作评估对整个社会的影响。随后，在1989年3月，巨型油轮“埃克森·瓦尔迪兹”号在阿拉斯加州威廉王子湾的布莱礁搁浅，原油泄漏量达到1100万加仑，造成了迄今为止美国历史上最严重的环境灾难。在该事件引起公愤后，美国国会通过了《石油污染法》，条件价值评估法也首次纳入了诉讼领域。美国国家海洋与大气管理局的任务是确定使损失货币化的方法，由诺贝尔奖得主、经济学家肯尼斯·阿罗（Kenneth Arrow）和罗伯特·索洛（Robert Solow）领导的蓝绶带专家小组审查了条件价值评估法的优缺点。美国石油和矿业部门支持的大多数商业协会都反对条件价值评估法的合法性，他们认为这种“主观”方法只能用于支持广泛性监管和大额损害赔偿。令他们惊讶的是，事实并非如此。蓝绶带小组对条件价值评估法做出了不冷不热的认可，对其未来的使用设定了许多参数。此外，顾问之间的意见差异导致政府机构对“埃克森·瓦尔迪兹”号的诉讼案只能庭外解决，得出的赔偿金额为11.5亿美元（而平均估计的损失总额达到了近30亿美元）。埃克森公司自身资助了几项“反向研究”，最初是为了质疑原告委托估价的可靠性，后来则影响了专家使用的方法和技术。自那以后，在生态损害评估中应用条件价值评估法的研究都受到了石油工业的很大影响。

不用说，条件价值评估法引起的争议不仅仅因为其潜在的政治化倾向，还因为其明显的概念性问题。如果市场研究人员承认估计传统市场商品支付意愿的本质性困难，那么毫无疑问，在调查不熟悉的商品和概念（如生态系统服务的总经济价值）时，问题会进一步复杂化。通过“总经济价值”概念，经济学家得以理解使用价值、非使用（或存在）价值和选择权价值的结合。使用价值是经济学中的一个经典概念，它表示人们使用特定商品时获得的效用（通常是为了实现其他特定目标，但也不完全是这个目的）。一把锤子的价值就相当于我实际上能用它来做什么。作为一种工具性概念，使用价值对木材、鱼、石油、水等自然资源的估值非常重要，因为它们可以被转化为建筑材料、食品和能源。然而，自然的价值不仅仅体现在它特定的用途上，它也有其内在的非使用价值。从欣赏日落或壮观的景色得到的愉悦感不是因为人们能够使用它，而仅仅是因为它的存在（这也就是为什么非使用价值也被称为存在价值）。最后，大自然是具有多种应用的实用源，其一部分作用目前可能还不为我们所理解或知晓。所以，它的整体价值不仅包括了使用性功能和非使用性功能，还包括它的潜在用途。这就是成本效益分析

师所称的选择权价值（或期权价值），表示通过对资源的保存，未来可能会获得的效用。

这些不同（且互补）的价值定义不仅在概念性角度相关。它们也有重要的政治（和法律）性牵连。以环境损害诉讼案件为例，尽管算作使用价值的损害仅能由因特定生态系统退化而造成直接经济损失的人进行索赔（例如降水量减少给农民带来损失，水污染给渔民带来损失，空气质量差给当地居民带来损失），但对于非使用价值和选择权价值来说，理论上每个人都可以要求赔偿（也被称为受到了被动损失）。不要在意衡量生态系统使用价值存在的明显困难，它对人类生活有着根本性但在很大程度上被忽略的影响。但由于主观性问题和认知（以及文化上的）限制，对非使用价值和选择权价值的估计充满了难以克服的复杂性，所以几乎不可能产生任何有意义的估值。例如，埃克森公司委托的一项研究对于法院应用条件价值评估法的可接受性提出了质疑，研究发现，在关于保护2000只、20000只或200000只鸟类的调查中，受访者给出了相同的估价答案，这种情况驳斥了受访者能将生态保护准确货币化的能力。相反，可能会有人争辩道，受访者只是在应用他们的常识，他们会认为，无论保护多少鸟类，都算作是一种独立的价值。也许他们认为鸟类的生存是无法定价的，也可能他们根本对此毫无概念。科斯坦萨及其同事认为，受访者“可能并不了解情况”，他们的偏好“可能无法充分体现社会公正、生态可持续性和其他重要目标”。换一种说法：

假如我们真的生活在一个生态可持续、社会公正的世界里，而且在这个世界里每个人对他们与生态系统服务的联系都有很好的理解，那么市场价格和支付意愿调查都可能产生与现在完全不同的结果，生态系统服务的价值可能会增加。

在一个短期效用、自身利益和消费被提升为结构性行为准则的社会中，少数实验参与者或调查对象（更不用说所谓的专家焦点小组）可以得出世界自然资本的可靠估值，这一点是值得怀疑的。而且，所有这些方法都理所当然地被视为以人类为中心的评估观点（自然的价值完全由人类能从中获得什么体现），完全无视环境具有内在价值的可能性，这是伦理学家和自然科学家经常提出的问题，还忽略了以生态系统为中心和以生态为中心的观点对自然更广泛的影响。正如我们在第三章中讨论到的那样，成本效益分析和经济评估是对问题危险地简单化，很容易误导决策者。即使那些接受货币估值有效性的人也需要认识到，“这些价值存在太多的经验不确定性”，它们“反映了我们对物质世界的有限理解。”事实上，任

何自然科学家都会认同，人类并不知道生态系统提供服务的所有方式，以及这些方式如何随时间变化。评估人员也改变了他们对生态系统作用的看法，因为他们的信息基础发生了变化，社会背景也随着时间发生了变化。

在很多方面，1997年《自然》杂志上的文章和2002年《科学》杂志上的后续研究成果都更像是“政治宣言”，而非科学测量。这两项研究都旨在“激发更多的研究和争论”，几十年来，大自然的估价一直以狭隘的个人效用最大化为主，而这两项研究帮助解决了这个问题。这些学者意识到他们最初的自然资本估值方法存在矛盾和局限性，他们承诺通过探索新方法来捕捉两个重要的额外目标，即生态可持续性和社会分配公平性，因为他们意识到“当评估仅仅基于当前个人偏好和效用最大化时，就像在传统分析中做的那样，并不一定会形成生态可持续性 or 社会公平”。一项具备公平性的评估要求个人在选择偏好时将自己视为团体的成员，而非单独的个人。然而，这就需要一个非常不一样的过程，其中要包括整个社会的公开辩论和协商，以达成“对当前和未来社会（包括非人类物种）所有成员公平的价值共识”。由于评估过程中总是存在赢家和输家，只有当参与者蒙上“无知的面纱”，集体辩论才会有成果，就像约翰·罗尔斯的正义论一样，每个人在投票时“都对他们目前或未来社会中的个人身份一无所知”。最后，基于可持续性的评估需要全面地评价生态系统服务与全球体系的物理、化学和生物功能的关联方式。如果我们认为所有物种，无论从人类中心角度来看是多么无趣或缺乏直接效用，都可以在自然生态系统中发挥根本作用，那么生态系统的估值就不应该直接引用人类的偏好。长久以来，人类在生态系统中的运作“就好像他们是整个系统的代表一样”。相比之下，真正的可持续发展方法则需要关注对“相关生态经济系统的贡献”以及“为后代选择的机会”的整体评估。不用说，达成这样的目标需要使用完全不同的方法。而且，由于“涉及大量不确定因素”，任何模型在使用中都需要加以预防，“寻找可能的价值范围，小心谨慎避免犯错”。

## 自然界的金融化

但是，我们就能这样定论了吗？我们有没有对新模型进行必要的细心对待，试图将对价值的整体理解整合起来，并始终谨慎对待它？恰恰相反。就像减缓气候变化领域蓬勃发展的补偿行业一样，自然资本核算已经成为私人顾问公司、银行和投资集团的新业务，这主要得益于全球治理的新形式。（本书分享更多索搜@雅书）

2006年，联合国环境规划署委托进行了一项研究：估算加勒比海和印度洋的珊瑚礁价值。这两片海域是潜水者重要的旅游景点和渔业的重要来源，为当地经济创造收益。结果就是：珊瑚礁的价值每年在10万~60万美元，红树林的价值则在20万~90万美元。在这些生态系统有助于维护白色沙滩的地区（另一个重要的旅游景点），估计每年的价值可以高达100万美元。同年进行的另一项研究认为，各种昆虫（尤其是数量正在下降的蜜蜂）每年为美国经济贡献了至少570亿美元。

2009年，《经济学人》杂志报道了圭亚那中部占地37万公顷的热带雨林伊沃克拉马（Iwokrama）的故事，说该热带雨林刚刚进入了“全球经济系统”。当热带雨林的受托委员会认识到捐助资金不足以维持森林，他们便将其推入了市场。首先，雨林引入了木材开采、生态旅游计划等获利计划，将森林产品进行商业化，比如出售蜂蜜、植物油，在雨林进行生物勘探和林业研究等。然后，它决定出售用于测量和评估森林生态系统服务的许可证，该许可证由伦敦的一家投资公司树冠资本收购。自那时起，后者便开始营销“生态系统服务证书”，附有十年期可交易债券，将碳上涨成本和再造林项目金融化作为补偿机制REDD（减少毁林和森林退化造成的排放）和REDD+（针对发展中国家的REDD机制）的一部分，这是《京都议定书》引入的两个全球机制。正如树冠资本宣称的那样，那些通过将自然资产货币化来保护森林的国家很可能会将这些资产转化为每年价值数十亿美元的全球资产：“投资界开始意识到这个机会”。科学家和投资者联合经营的全球树冠项目（The Global Canopy Project）一直引领着全球对热带森林货币化的探求，也为此出版了流行书籍，如《森林金融小手册》和《生物多样性金融小手册》，这些小册子解释了经济回报在自然资本领域看似无止境的可能性。也许这并不奇怪，该项目的受托人董事会是由投资银行家和金融家主导的。2011年，树冠资本

举办了一个名为“解锁森林债券”的研讨会，该研讨会得到了世界最大投资银行高盛的资助和支持。与会者一致认为，“发行债券直接解决了时间和规模的问题，使得发行人可以现在就筹集大规模的资金，这些资金会由现有的和预期的未来收入进行偿还。”尽管如此，他们强调了债券发行人在说服投资者时面临的挑战，尤其是当如今碳市场变得不太可靠时，要说服投资者相信“计划偿付债券的现金流是足够安全和可预测的”显得有些困难。因此，为了使森林看起来有利可图，他们建议引入“生态系统服务市场（如水、生物多样性）、可持续木材和农业市场、监管（如税收）和森林友好型贷款（例如贷款给依赖生态系统的中小型企业）”，这需从政府得到必要的支持，“以确保这些现金流量的实现，使森林保护成为一项具有吸引力的投资。”

“2010绿色发展机制倡议”，最近更名为“绿色发展倡议（Green Development Initiative）”，是一项公私合作伙伴关系，在联合国《生物多样性公约》的支持下用以发展“创新的市场金融机制”。其主要目标是根据认证标准和独立审计指导生物多样性管理中的私人投资，它需要对生物多样性的经济价值进行评估。正如专业性杂志《生态系统市场》（Ecosystem Marketplace）认为的那样，认证的农业和森林产品、私人土地信托、流域和其他环境服务的付费以及生物多样性丧失的补偿可能很快就会比肩清洁发展机制下补偿计划产生的数十亿美元。在澳大利亚，生物多样性银行（也被称为“生物银行”）已经被引入，以保证承诺增强和保护生物多样性价值的土地所有者所获得的生物多样性信用额度，这种额度可以出售给开发商以平衡（或补偿）新建筑项目可能带来的对生物多样性价值的影响。2011年，包括世界上最大的两家消费品公司雀巢和联合利华在内的一些跨国企业加入了一个《自然资本领导契约》，据称这是为了刺激企业对自然产生积极影响。在2012年6月的里约+20峰会上，二十多家跨国公司表示，要通过各种方法来评估经营和投资过程中自然资本的价值，其中包括可口可乐公司、陶氏化学公司和耐克公司等。在他们题为《新业务势在必行：评估自然资本》的报告中，他们确定了从这项举措中可以获得的一系列好处。在同一次峰会中，一些世界领先的私人金融机构签署了《自然资本宣言》，其中承诺将环境核算纳入其投资计划和经营中。用他们的话说，“每一种经济活动都会对自然资本产生影响……这些影响可能会导致重大财务风险，但也会产生相关的商业机会。”

虽然这些新的治理趋势被视为重要突破，但应该以怀疑态度来看待私人金融集团和企业巨头对自然资本核算和生态系统经济越来越多的介入。对大部分宣称



的“绿色”计划的诚实性也要进行慎重考查。首先，由于高盛银行参与到引发全球金融危机的事件中，更具体地说，由于它在希腊主权债务危机中扮演了可疑角色，帮助该国领导层向欧洲当局掩盖了其实际财务状况，高盛在森林债券新业务中的参与引起了高度关注。参与这些新“交易”的大多数其他商业和投资银行也深深卷入了导致欧元危机的金融产品的市场化。《自然资本宣言》的签署者之一，意大利银行业巨头裕信银行（UniCredit）的业务范围覆盖20个国家，意大利当局以其逃税为由起诉了该银行，且控告该银行将“有毒衍生品”合约主动转移给投资者和消费者以清理账户。生态系统和生物多样性经济学（TEEB）计划背后的策划者也是德意志银行环球市场部门的主管，德意志银行已经成为欧洲银行业精英集团的标志，也是欧洲债务抵押债券的主要驱动因素之一，它已经在伦敦银行间同业拆借利率（Libor）和欧元银行同业拆借利率（Euribor）丑闻中因为涉嫌欺诈、间谍活动、逃税和操纵利率调整接受了调查。在该领域的一些引领参与的公司也存在相似的情况。例如，像雀巢这样的公司，在食品生产和销售方面的记录也有很多疑点。自20世纪70年代以来，全球的民间社会团体一直在引导对雀巢产品的国际性抵制，因为他们声称该公司在非洲国家进行了强制和不道德的营销，谴责雀巢的母乳代用品（婴儿配方奶粉）导致了非洲儿童的营养不良和死亡。雀巢公司也因为拒绝标注转基因食品遭到了卫生组织的批评，因为从破坏雨林的子公司购买棕榈油遭到了非政府环保组织的批评，因为与缅甸军政府有生意往来遭到了人权捍卫者的批评，以及因为剥削农民和侵犯工人权利遭到了工会的批评。在一部2005年的电影纪录片中，当时雀巢公司的首席执行官彼得·包必达（Peter Brabeck-Letmathe）对有机农业进行了抨击，他为转基因食品的大规模商业化辩护，并直率地指责了非政府组织在捍卫水资源作为公共权利时的极端做法。可口可乐公司在环保领域也无法拥有一份闪光的“履历”。除了在南美因为打击工会备受谴责，这家软饮料巨头还在印度等国建立了寡头垄断的水资源分配控制体系，当地农民指责其设备消耗了他们的可用水资源，联邦政府和地方政府也已经对其大规模杀虫剂的使用展开了调查。陶氏化学公司不仅是美国军队在越南使用的橙剂<sup>[3]</sup>的主要生产商之一，而且还始终拒绝清理1984年世界上最严重的工业灾难——博帕尔毒气泄漏案造成的污染。这场灾难发生于美国联合碳化物公司属下的一家印度农药厂，陶氏化学于2001年收购了该公司。

全球咨询公司麦肯锡除了与世界上一些最大最强的公司（被《财富》杂志列入全球最具领导力十大公司也令他们引以为傲）合作外，也是为政府在造林和再造林政策上提供咨询服务的先导者。造林和再造林政策作为REDD+计划的一部



分，近来一直被卷入一些最引人注目的金融丑闻。能源巨头安然公司的首席执行官杰夫·斯基林（Jeff Skilling）因犯欺诈和其他联邦重罪而服刑14年（由24年的监禁减刑），安然曾经是麦肯锡的重要合作伙伴。据《华尔街日报》报道，麦肯锡多年来一直是安然公司的战略顾问，由于安然公司在财务和运营上存在多年舞弊问题，其破产结果（当时是美国有史以来最大的公司破产案）也引发了对咨询公司的责任问题的争论。后来，麦肯锡公司卷入了有史以来最大的内幕交易案——帆船集团（Galleon）丑闻。检察官称，麦肯锡公司的一些高级管理人员向帆船集团对冲基金泄漏了麦肯锡客户的内幕信息来非法获利，其中包括麦肯锡的高级合伙人兼董事阿尼尔·库马尔（Anil Kumar），他对共谋和证券欺诈指控表示认罪，并揭露了麦肯锡前首席执行官拉雅·古普塔（Rajat Gupta）在其中的牵连，后者在2012年被判处了两年监禁。

在环境治理方面，由于全球温室气体减排成本曲线的广泛采用，麦肯锡已经在该领域拥有了显著地位。这项于2007年设想的成本效益分析概述了逐步减少排放的不同选择，曲线高度代表成本，每部分的宽度表示碳减排的相对量。成本曲线在政策制定者中越来越受欢迎，尤其是因为它更加强调节约成本（花费成本少）的估量方式。但绿色和平环保组织认为，麦肯锡的方法过分简化了事实，且存在着对比较成本不切实际的假设：

如果替代当地自给自足农业的真实成本被低估……比方说忽视了交易成本和更广泛的社会和环境影响，而处理工业采伐的成本被高估（例如夸大伐木的经济价值），并且这些假设被置于成本曲线上，那么基于曲线做出的每一个决策都会更有利于伐木者而不是小型农业生产者。这种结果不仅会造成社会性破坏，而且可能无法实施，在经济上的不合理会导致无法有效减排。

同样令绿色和平组织感到遗憾的是，虽然麦肯锡声称其分析“依靠事实”，但却拒绝透露其计算中依据的数据和模型，并以公司的知识产权为由进行辩解：“外部世界无法知道麦肯锡是如何估计各种减排措施的不同成本的。”

在这种情况下，我们并不奇怪会在英国自然资本委员会发布的常见问题中，发现这样的疑问：“我们能否相信会计师和经济学家能得出自然的真正价值？”虽然委员会已经竭尽全力解释评估结果是由社会科学家和自然科学家组成的多学科评估组做出的，但金融与企业权力之间的明显联系，以及旨在衡量自然价值的大多数举措，都仍然令人担忧。正如记者克里斯汀·麦克唐纳（Christine

MacDonald)在她的书《绿色公司：环保组织内部人士揭露环保事业是如何衰落的》中所说的那样，大多数这些政策只是为某些在环保上存在过失和无视基本权利的臭名昭著的公司提供了一种绿色装饰。

在这样的背景下，绿色发展机制的发起受到全球环境运动的批评并不令人惊讶。该机制将应用《京都议定书》中清洁发展机制的理论基础来支持生物多样性，但许多人怀疑私人投资者支持生物多样性动机的真诚性，他们认为如果在支持生物多样性上能获得切实的公司利益，“起初就不会存在生物多样性丧失的问题”。名誉、财富和权力被视为对“环境事业”产生新兴趣背后的根本动力。其他人，特别是在工业化程度较低的国家，则警告说，生物多样性的金融化最终将导致对当地人口财产剥夺、驱逐和暴力的全球性行为，其中包括原住民运动，而他们才是生物多样性真正的守护者。像生态旅游这样表面无害的概念用来在许多地方区域刺激消费欲望，并引入了非当地的生活方式和议程，它们会一直影响原始生态系统的传统均衡。几个世纪以来持续为社区供应食物的自给生产体系也受到了企业因对生物多样性金融化的兴趣所带来的负面影响。

社会学家迈克尔·戈德曼（Michael Goldman）在他的书《帝国自然》中指出：

在不同寻常的同步性中，可持续发展人群和新自由主义发展人群联合起来，在南方重建自然，将大片由社区管理的非资本化土地转变为用于商业采伐、药物的生物勘探、出口经济作物、动物保护和特色生态旅游的跨国管制区。

由此，生物多样性保护已经成为大规模“绿色掠夺”浪潮的托词，有大量土地被用于进行生物多样性保护、生物碳封存、生物燃料、生态系统服务、生态旅游和排放补偿。这种环境保护的全球热潮由一群新行为者和特别的联盟牵头，其中包括“养老金和风险资本家、商品交易者和顾问、地理信息服务提供商和商业企业、生态旅游公司和军队、环保人士和焦虑的消费者”。一些国际公约、体制改革和治理协议增加了其合法性，它们宣称其目的是拯救自然和应对气候变化，但最终可能导致全球性的共享空间“封闭”。

环境学者布拉姆·比舍尔（Bram Büscher）、沙恩·沙利文（Sian Sullivan）和他们的同事将企业和金融界的这些趋势描述为“新自由主义生物多样性保护”的一个例子，他们将其解读为“意识形态和技术的融合，其前提是只有屈服于资本，在

资本条件中进行再评估，自然才能被‘保存’”。在他们看来，这种霸权话语对于当前的资本制度至关重要，因为它打开了投资机会和新的积累领域，同时“巩固了普遍共识的表象”，这种共识实际上建立在非常有争议的价值观和方法上。通过数据、价格、成本曲线和其他衍生工具，这一金融化过程“能将其自身置身于它激发的矛盾领域之外，即使在批评和歪曲这些矛盾时也是如此”。

## 结论：自然有限公司的成立

2008年，英国广播公司推出了一部名为《自然有限公司》的系列纪录片，承诺向大家展示“我们在多大程度上依靠着自然才能保持经济进展”。联合国环境计划署对此表示了部分支持，该系列纪录片中摄制了科学家、生物多样性业务的投资者以及金融专家。没有其他的标题能够更好地总结过去十年来发生的自然资本市场化进程，它被置于公众监督和民间社会辩论的“雷达”之上。新核算模式、条件价值评估法和其他形式自然资本货币化的试验最终导致了金融和公司利益之间，以及整个自然生态系统之间的危险勾结。正如上一章讨论的气候变化减缓一样，经济推理和数值模型的使用有助于在对自然进行估值的过程中形成一种中立的光环。数据已被用于赋予抽象概念实际的权威，比如成本、收益、价值和价格。自然货币化的过程依赖于数据提供客观性的能力，而实际上，正在做出的决策仍然是主观的。

大多数倡议的最终目标可能是真实的。它们可能确实会诚实地应用经济模型来保护社会和环境质量。它们认为，定价会迫使人们进入“理性的决策思维框架”来分析某种发展轨迹的“收益与损失”。运动服装巨头彪马是第一家披露“环境损益账户”的公司，它率先组建了一个愿意将环境成本纳入价格标签的企业联盟。彪马老板表示，“通过以欧元和美元显示环境成本，我们新推出的环境损益

（EP&L）使彪马产品给环境带来的影响变得具体化，并使得每个人都可以轻松比较产品的可持续性。”2014年，数据机构特鲁克斯特公共有限公司（Trucost）推出了自然资本领袖（Natural Capital Leaders）指数，以评估商业对自然资本保护的承诺。在将条件价值评估法应用到计算环境损害成本的过程中，石油和矿业行业公司最初的阻挠证明了自然资本的“主观”估值也可能损害公司的利益。所以这种工具可以说是有利也有弊。

全球林冠项目计划认为，市场计划可以帮助集中资源，防止投机力量对生物多样性的威胁，特别是在农产品方面：“我们需要揭开神秘面纱，探索创新融资机制，这可能会有助于刺激新的自然资本经济。”如果国际市场通过对牛肉、大豆和棕榈油等商品的支持破坏了热带森林，那么市场同样可能产生强大但相反的反应来阻止市场对大自然的破坏。对他们来说，市场可以用来对抗市场。拯救大自然的唯一方法就是将其出售。

毫无疑问，国际金融和企业力量能够利用环境危机获得新的社会认可。他们通过宣称拯救环境，为自身寻求新的合法化，特别是在2008年金融危机之后，这种诉求更加明显。他们的方法已经被“环境话语的政治学”所强化，集中于理性、可持续性、管理主义和现代化等概念，并基于成本效益和效用最大化的原则。推行保护未来经济利益潜力的策略的逻辑延伸到了人类的行为动机主要是个人利益的假设上，最后的结果就是导致财富和幸福的集聚。在对世界的这种设想中，经济模型可以将盈利能力与健康的生态系统结合起来，这看起来像是最成功的双赢局面。

市场模式（无论是破坏还是修复自然）是建立在特定的人类学范式之上的，通常被定义为经济人。经济人的偏好是被赋予的（也就是说，它们不是在文化和社会上产生的），而成本效益计算管理着他的自利理性。在经济人的现实中，市场无处不在。如果某些资源或服务没有市场，那么就可以通过实验和问卷调查“模拟一个伪市场”。正如我们所看到的，成本、价格和收益的数据化呈现过度简化了现实生活过程的复杂性。将生态系统简化为有价格的商品会使自然受到人为经济规则的管制。更重要的是，它掩盖了环境治理中总是有输家和赢家的事实（这是经济人不愿意接受的东西），从而精简了分配问题。成本的概念是误导性的，它基于价格，而价格是社会中的人工产物。价格并不存在于现实中。真正存在的是基于多种因素将名义价值归于产品的意愿，这些因素包括稀缺性和吸引力，以及传统、习俗和其他类型的社会影响因素。所有这些因素都是暂时的：根据定义，它们会随着时间变化。根据这种主观和临时结构来构建整个治理体系，再好也只能说是短视的。

有趣的是，自然科学家通常对依据数据处理问题感到自豪，他们不愿意将高度抽象的模型应用于环境治理领域。生物学家戴维·埃伦费尔德（David Ehrenfeld）是保护生物学会官方杂志的创始编辑，他对此表示：

将所有保护问题简化至经济角度都是危险的，且会产生相反的效果。相信市场力量和供求规律能纠正不公平、恢复健康均衡在经济学中是行不通的，在保护方面肯定也不起作用。

为了解释社会建构主义，我们可以将成本描述为“我们希望某些事物成为的样子”，由此我们最终决定了成本是什么。是这样吗？由于我希望的贴现率与您所想不同，所以在自然的估价中，究竟使用了谁的贴现率，这点尤其容易让人怀疑。

而答案通常是：这一利率来自市场。这就意味着，金融市场中的预期和预测会成为指导自然金融化的参数。我们并不介意衡量自然资本对经济增长的贡献用的是由联合国发起的采用GDP框架的SEEA，还是用的石油行业顾问采取的条件价值评估法，抑或用的衡量生态系统服务的支付意愿的市场模型。

正如前一章中讨论的那样，在成本效益分析方面（特别是在环境领域）廉价的方案对整个社会来说，往往会变得更昂贵。所有基于成本的运作都只能用以测量有限的维度。定价过程非常具有选择性，往往忽视基本要素而仅仅因为它们不能在市场条件下被估计。无法定价的东西便不存在花费，因此也没有进入成本效益分析的等式。但不论它是否有价格，最终都要有人为其“买单”。这些最后的付款人，往往是我们。一些情况下，是现在的我们；另一些情况下，则是我们未来的后代。

赋予生态系统价格可能会迫使我们认识到自然界的经济贡献。然而，它也可能带来自然资源商品化的危险，因为任何有价格的东西都可以买卖。正如本章中讨论的，大多数人可能并不能迅速掌握使用价值、非使用价值和选择权价值的复杂联系。当我们看到价格时，我们想到的是另一种价值：交换价值。事实上，有价的东西可以出售，或者换成具有相同价值的东西。价格创造了一种幻象，即自然产品和服务可以在市场上交换，就好像它们是传统的生产要素一样。自然科学中有一种“强可持续性”的概念，也就是说，某些资源稀缺且不可替代，因此人类活动必须受制于环境资源的有限性。这种概念现在倾向于被“弱可持续性”的经济原则所取代，“弱可持续性”原则认为任何类型的资本都可以作为生产投入，完全替代自然资本。这种逻辑反复想表达的意思是人类可以征服自然并最终实现自然的自我拯救。因此，投资基金机构气候变化资本集团（Climate Change Capital）认为，如果在一个国家用尽了现金（由于对私人银行的持续金融救助等因素）的世界里，“我们需要转变为自然保护募集资金的方式，而且这一转变需要我们迅速增加从私人部门获得的资金”。我们需要一种新的“货币”，它需要与生物多样性信用等价，或有“兑换率”，“这对通过建立更深入和更流动的市场以吸引投资至关重要”。总之，解决环境问题的唯一办法是将自然置于私人投资市场的控制、语言体系和管辖之下。

生物学家知道，生态系统是复杂、多方面均衡和持续演变的结果，在这个系统中所有部分都结合在一起。对于自然科学家来说，构成地球的互联网络是一个

各部分相互依赖的万花筒，每个部分对整体都是至关重要。没有自然，就不存在人类。而相比之下，自然的金融化将生态系统分裂、分离、碎片化并最终使各种“可销售的”部分通过衡量、估价和交换各自疏远。就像在金融业一样，“自然衍生品”市场必须被切割开来，以适应投资者的需求和他们对财务回报的期望。

这种推理产生了明显的矛盾。2000年，联合国秘书长呼吁进行“千年生态系统评估”，为自然资本及其与人类的关系提供最高水准的分析。这项调查的结果于2005年发布，有超过1300名科学家参与编写了报告和估计结果。他们得出的结论是，人类活动已经使“地球处于物种大灭绝的边缘”，“除非人类的态度和行为发生改变，否则未来几十年全球生态系统的压力将会持续增加”。但在报告中，科学家也发现，尽管生态系统服务出现了这种崩溃的状况，但人类福祉确实有所增加。这一发现与数十年的环保运动相违背，因为环保活动者认为，生态退化最终会导致人类福祉的下降。后来被称为环保主义者悖论的观点似乎证实，技术和现代性最终已经将人类福祉从自然过程中分离出来：这种悖论表明，更多福祉可以成功与自然损失交换。

针对日益激烈的生态系统服务货币化论战，环保组织绿色和平指出，为“互相紧密联系的自然系统”赋予数值“本质上是投机的，且不一定明智”。将自然简化为粗略的数据可能具有危险的误导性，因为它忽略了“自然系统的相互关联性”和“临界点与突变的可能性”。此外，它给了我们这样一种印象，即人类可以把自然控制为“资产”，在地球系统崩溃时，人类可以“拯救”它。但地球并不是一个金融市场，如果我们忽略地球的承载限度，“救助可能会变得太迟，世界上不会有任何金钱能帮助我们”。

正如本章所描述的那样，若统计数据会为了服务不同的利益而被利用或者滥用，一项好的事业也非常容易变成极其危险的事情。拯救我们的生态系统是一项值得赞扬的使命，在一个痴迷于数据的世界中，某种形式的衡量会不可避免地用以估测保护政策的规模和范围。但是，当这些衡量变成价格时，就像在自然资本和生态系统服务核算领域经常做的那样，市场会挤掉其他形式的治理。通过一系列表面中立的方法，无法估价的东西被评估、定价，然后变成了商品。将自然的复杂性转化为数据的简单性，为狭义的经济方法铺平了道路，它将自然视为必须产生经济回报的投资。为了计算，自然必须成为人类的所有物，并被赋予“生产力”。在危机发生的世界里，金融市场需要扩展到新的领域以创造利润，自然的金

融化很可能为资本积累和投机开辟新的可能性，这会以最大限度损害我们的生态系统和社会。

正如约翰·梅纳德·凯恩斯（John Maynard Keynes）1933年在都柏林大学题为《自给自足的国民经济》的演讲中所说的那样：

我们摧毁了乡村的美丽，因为大自然未被占用的光彩壮丽不具有经济价值。我们能使太阳和星辰的光辉黯然熄灭，因为它们不产生股利……但如果我们不再屈服于对会计利润的追求，我们就可以改变文明的发展轨迹。

注解：

[1] 卫星账户是以国民经济核算体系中心框架为基础，对其某些概念加以修改而建立的附属核算体系。运用卫星账户形式，把中心框架的有关内容直接或间接地引入其中，再加上补充的或替代的概念，既避免了对中心框架严谨统一性的破坏，又极大地扩展了国民经济核算体系的分析功能，为满足不同需要提供途径。如旅游卫星账户、资源环境卫星账户等。

[2] 在这种拍卖中，竞买者同样以密封的形式独立出价，商品也出售给出价最高的投标者。但是，获胜者支付的是所有投标价格中的第二高价，所以它被称为第二价格密封拍卖。

[3] 橙剂是一种高效落叶剂，因其容器的标志条纹为橙色而得名。



---

## 第五章

# 数据能够为人类带来利好吗？

---

“并不是所有重要的事物都可以被计算，也不是所有能够被计算的事物都重要。”

——阿尔伯特·爱因斯坦

“‘发展说’的力量来自于它的引诱力，就‘引诱’一词的各种含义而言，即：吸引、取悦、诱惑、致幻，同时滥用力量、背离事实以及欺骗。”

——吉尔伯特·里斯特

《发展的历史：从西方世界起源到全球信仰》，1997年

**测**量方法帮助我们度量是否以及在怎样的程度上实现了预设成果。医生依赖数据来检测患者的治疗效果；工程师计算出数据来评估建筑物的稳定程度；机械师利用数据来测量车辆中水、油和燃料的体积。正如我们看到的，数据本身不具有标准价值，它们的力量来源于其将复杂现象简化为少数可观察事实的能力，这就是为什么，数据推理在被系统应用于人类世界时，会引起各种各样的误差，而这一现象在发展合作与社会变革的政策中体现得尤为明显，因为在全球的项目中，数据测量是设计与评估的关键。关于国际援助有效性的辩论基本囿于对于验证发展政策效果的各种量化指标的确定上。在社会变革领域，对影响的评估很大程度上是由计量经济模型决定的，而在这些模型当中，数学算法的变换导致社会关系的复杂性被忽略。在气候变化与自然资本核算的议题当中，这些成功的测量方法已经为科技工具的传播开辟了道路，从商业到慈善业，再到非营利行业。在传统的以定性分析、长远角度及社会参与为特点的部门当中，诸如社会投资评级、社会投资回报、成本效益以及标准评估等概念都变得格外流行。

作为一名年轻学者，我自己也经历过这样的变革。在2005年左右，我受邀为欧洲最大的发展组织之一提供建议，其名称不便透露。那时我在为非政府组织和

政府就建立社会指标与开发参与性评估工具提供咨询。我记得在九月美丽的一天，我的客户对着一栋引人注目的多层办公楼夸夸其谈，那栋办公楼的主入口直通市中心的观光街，周围坐落着大多数国际发展机构的总部，办公楼的走廊里挂着各式各样的照片：非洲的贫困儿童、自然灾害和融化的冰川，那让人感到仿佛全球问题的解决都能够在这里毕于一役，是充满力量又令人哀恸的一幕。

他们邀请我去优化影响评估工具，而我一直对影响评估持怀疑态度，在我看来——正如任何有一点常识的理性的人看来，在现实社会现象当中追踪因果过程是一项十分艰难的任务，由于概念与方法的复杂性，这项任务几乎难以实现。同时我也意识到，仅仅通过拒绝引入影响评估工具来防止“影响”问题也同样天真。经验的确显示，当没有方法论可以用于测量影响时，各种各样天花乱坠的观点便会应运而生。这是一个令人头疼的困境：测量影响难度极大，通常不可能，但无所作为又意味着市场策略（比如公关活动、煽情的图片和其他形式的无稽言论）会在缺乏明显的成功证据情况下乘虚而入，吸引公众眼球。这也是我接受那份工作的原因，我诚心致力于寻找充满发展前景的新测量工具，以推动影响评估工作进步。

那个时候，大多数传统捐赠者，比如经济合作与发展组织中的政府与传统的慈善机构，都经历着普遍的发展“疲劳”。他们在发展中国家工作了几十年后，发现发展援助的投入实在难以维持，不公正、不平等与地方性贫困大行其道，援助对象国的情况鲜有起色，而新的打击总是接踵而至，这些对象国似乎困于腐败、社会失衡的恶性循环中。而且，右翼党派在欧洲的选举中大获全胜导致了普遍的“抨击援助”的公共争论，越来越多的人开始质疑：发展援助使得国内财富流失，继续海外投资意义何在？公共预算被削减、开支审查真正妨害着社会福利体系。与此同时，失业与仇外现象屡见不鲜。如果发展援助事业想要生存，那么它们需要有力地证明自己的价值。

尽管意识到了潜在的挑战，我还是接受了那份工作，并且希望我的客户能够打破陈规，这也是为什么我进行了一场关于多行业分析与自下而上参与需求的报告。我强调，参与方法非常重要，它可以通过利用网络工具被进一步推广，同时我也希望，在自上而下的传统评估从来都行不通的领域内利用大众合作的潜力。我希望他们能够意识到，信息交换与开放的参与过程本可以减少影响评估的成本，同时可以对全球大量的受资助项目的直接/间接影响进行追踪。而且，我表

示，经过整合的影响评估参与过程不应被视为原本紧张的预算与超负荷发展事业的负担，相反，整合的评估系统应该被视为发展工作的重要组成部分。良好的、持续的影响评估过程应当强化发展项目本身，甚至被证明能够最有效地产生实质性影响，因为参与者的授权是任何一种社会、经济发展的前提。当我总结陈词的时候，他们注视着我，满脸狐疑。显然，我参与性的、自下而上的提议并没有给他们留下深刻的印象，他们也并未理解我所说的奇妙的大众合作。接下来的几秒钟里，会议室一片寂静，然后，首席执行官——一位五十多岁的气质非凡的女士，抓起麦克风，斩钉截铁地对我说：“费尔拉蒙蒂先生，您一定是误会了，我们对网络和整合参与评估并不感兴趣，我们希望您能够算出一项数据来告诉我们，我们所做的到底有没有用，就这么简单。”

## 援助效力的政治：历史回顾

在我们生活的世界里，所谓的“富国”每年在援助低收入国家上都会花费几十亿美元，而在传统看法中（并且我相信这是正确的）所谓的发展中国家接受的最小国际投资额是17000亿美元，几十亿美元无异于一毛之于九牛。这些各种各样的发展资金，包括官方发展援助项目（比如由政府支付的政府开发援助）、公共和私人借贷、由其他机构与非政府组织运营的项目、由流动工人的汇款与非经济合作与发展组织国家——这些国家还有专门的发展援助委员会——资助的援助计划支付的款项。尽管这些援助金额的总量只占全球经济产量的2.5%（2012年全球生产总值约为69万亿美元），但无论如何计量，这都是一个庞大的产业。

随着“二战”后布雷顿森林会议的召开与国际复兴开发银行为重建欧洲做出努力，国际援助体系在第二次世界大战的废墟中诞生了。如今国际复兴开发银行已经加入了世界银行集团与国际货币基金组织（IMF）。欧洲复兴计划——通常所说的马歇尔计划，建立于1948年，正式运营到1952年，引领了经济合作署——美国国际开发署（USAID）的前身的建立。1942年，哈里·杜鲁门（Harry Truman）总统发表了祝贺经济合作署成立的演讲，这一演讲标志着历史上“发展时代”的开端，并且也通过经济、军事（几个月后，这一方面的合作因北大西洋公约组织的诞生而被终止）合作重申了美国对联合国以及对重建欧洲的承诺。杜鲁门还提到了“第四点”：

第四，为了落后地区的发展与进步，我们必须着手一项大胆的新项目以促进科技、工业进步……这些地区的经济生活原始且僵化……有史以来第一次，人类得以利用知识与技术来减轻这些人的痛苦。

杜鲁门和他的顾问们都是新政的忠实信徒，他们也目睹了国内生产总值核算的发明不仅仅使美国赢得了冷战，也创造了举世无双的大众消费社会。对于他们，发展是经济技术的成果，这也是为什么杜鲁门在其演讲中平铺直叙地表示，此类为援助不发达地区做出的努力不仅会转化为资金，也会主要转化为科技援助。他解释道：“我们只能负担有限的用于援助他人的物质资源，但我们无法计量的科技知识却是持续增长、用之不竭的。”从演讲的几个段落中，我们能发现杜鲁门的第四点联系着发展模式的实质：将经济知识与市场基础上的治理方法从“发

达”国家输送到“发展中”国家。这一观点不仅深刻地影响着我们对于“发达”的理解——将非市场形式的经济体视为落后，也将对发展的追求仅仅局限于是一门技术：将关键经济原则（以市场为基础的原则）与机构移调至发展中国家。

根据杜鲁门演讲的第四点，联合国大会通过了“技术援助扩大方案（Expanded Programme of Technical Assistance）”，这一方案旨在将科技专家输送至发展中国家为管理人员提供培训，并且在后来发展成为联合国开发计划署（UNDP）。1951年，联合国发布了名为《欠发达国家经济发展应采取的措施》的报告（也叫《刘易斯报告》），该报告建议建立国际金融公司（International Finance Corporation），旨在为股票市场投资以及向私人企业借贷，这一公司于1956年由世界银行建立。在20世纪50年代早期，美国建立了国家发展援助体系，这一体系的建立得益于《共同安全法案》，该法案旨在为亚洲对象国家提供科技援助，正如其名称所表达的，它主要是为了筑起一堵安全的高墙，以抵抗远东地区共产主义的传播。

毫不意外，第一批接受美国援助的国家与地区包括韩国（以对抗朝鲜）、中国台湾（以对抗中国大陆）、伊朗（以控制欧洲）以及巴基斯坦（以对抗印度与中国）。1957年，后来的欧洲共同体也采取了相似的方法，通过建立欧洲发展基金来支持全球尤其是非洲曾经是欧洲殖民地的地区发展。这一时期，国际援助体系与冷战安全格局完全结合，并被用作杠杆以支持援助国的外交政策议程，尤其是对亚非拉的一些国家。

20世纪60年代，援助体系在制度上与财政上都变得更加复杂。1958年，世界基督教联合会向联合国大会会员国提交提案，第一次提出了发展援助与国内生产总值之间的最小比率：起初的目标比率是1%，但在后来的几十年内，该目标经修订会下降至0.7%。1960年，后来的欧洲经济合作组织（OEEC）发展成为了经济合作与发展组织（OECD），主要援助国也组成了开发援助集团（DAG，如今叫DAC），早期为欧洲国家（比利时、法国、德国、意大利、葡萄牙与英国）、北美国家（加拿大与美国）以及日本提供咨询。20世纪70年代，兴起的多边援助成为双边合作的补充形式，加上布雷顿森林体系的持续支持，国际发展援助已经从模糊的扶贫过渡为经济发展援助——这也被视作社会发展的最佳路径。援助政策与世界银行、国际货币基金组织促进低收入国家（自由）市场改革的多边贷款方案交织在一起。

吉尔伯特·李斯特在他的书中详尽地阐述了发展如何从一项“西方的创举成为全球信仰”，经济改革主义理论深刻影响了这一阶段的发展政策，尤其是沃尔特·罗斯托（Walt Rostow）的《经济成长的阶段》，该书成为解放“落后”国家的经典蓝图，因为书中的观点与美国强调的市场至上原则以及社会科技的发展方法相契合。罗斯托发现了五个经济发展的成长阶段：传统社会阶段、为经济起飞创造先行条件的阶段、经济起飞阶段、向成熟前进阶段以及高额大众消费阶段。罗斯托意图表达这样一个观点：经济发展轨迹类似于飞机的飞行轨迹，只有市场呈现最佳态势时，飞机才有可能起飞，社会发展就如同飞机沿着跑道加速，最后直冲云霄。在他的书中，罗斯托歌颂了“复利的强大之处”，这一现象使增长成为经济稳定的前提，尤其在财富基于信用之上的当今世界。在这一理论下，增长几乎被一种神圣的力量所维持着，这种力量确保“高额的大众消费变得普适”。罗斯托多此一举地开脱道：“这种对经济增长阶段的划分是一种主观、有局限性的看待现代史的方式，但在相对意义上，这一方式又是正确的。”罗斯托将自己的模式变成了法律般的权威。

这些年来，由于经济增长可以通过扩大原材料（如地下资源的煤与化石燃料）范围迅速得以实现，所以一些发展中国家很大程度上已经转变为出口导向型经济，并且这些国家（通常是集权国家）往往通过贷款来实行社会经济结构调整。尽管在20世纪80年代借贷国都未履行偿还贷款的义务，但越来越多的资金还是被借给了这些国家以弱化其民族工业和自由贸易的作用。随着冷战的结束，以往一贯对借贷国的民主责任制睁一只眼闭一只眼的政策也逐渐消弭，援助行业开始关注（或者说是纸上谈兵）更加“温和”的话题，比如人权、腐败和公民社会。虽然援助国的议程重点仍停留在发展上，但与良好的国家治理有关的因素也变得重要起来。援助行业再度转变其导向，却鲜有成功的例子，绝大多数尤其是非洲的发展中国家，似乎都受困于动荡与政治经济危机交加的恶性循环，同时地方性与其他形式的贫困问题仍旧猖獗。

在这一背景下，对于援助效力的争论开始在20世纪90年代后期出现。捐助机构、援助机构和私人基金会开始争相合作，援助对象国则力图证明援助的有效性以获得更多援助。2000年，联合国发布了“千年发展目标”，即八项于2015年实现的目标，其中具体包括：减轻极度贫困现象，将日生活费用少于1.25美元的人口减少一半、将儿童死亡率减少三分之二、将分娩死亡率减少四分之三，以及建立全球发展伙伴关系等。在国际政治中，“千年发展目标”成为促进发展的开创性转

折点，毫无疑问，这些目标巩固了数据在发展事业中的权威性，它们之所以获得广泛关注是因为其可测性与在目标数据上体现出的下降。但与此同时，联合国内部却哀声一片，他们认为“千年发展目标”严重脱离实际，并且，由于这些目标的数据结构，它们在发展问题上具有“一刀切”的弊端：错误引导援助的分配政策与“将一半以上的国家划入‘不合格’的分类中——因此最后削弱了政治家与大众对全球发展的支持”。联合国开发计划署被委托执行“千年发展目标”计划，并在其内部建立了新部门以实施基于成果的援助政策管理。而各级考核指标激增，经常削弱联合国开发计划署以往对能力发展的关注，而能力发展是一个培养技能、培育机构、建立关系的循序渐进的过程。许多人认为，对数据的关注尽管与联合国开发计划署的新重点——信用评级（见第二章）相契合，却代表了“该机构内部在倡导援助达到定性成果与通过定量成果实现援助效果的‘简化可测量’的分歧”。

“千年发展目标”也再次号召国际捐赠者给予对象国更多援助，而且重申0.7%的GDP目标（然而在2013年，只有五个国家——挪威、瑞典、丹麦、卢森堡与荷兰实现了这一目标）。2002年，国际发展融资会议于墨西哥蒙特雷召开，人们对援助效力提出了越来越迫切的要求。“伙伴关系”一词替代了传统的“关系”，并且，在2005年，援助组织签署了《巴黎有效援助宣言》（the Paris Declaration on Aid Effectiveness），敦促捐助者关注可测量的结果和切实的成果。

随着人们重新定义援助系统，关于援助效力的争论也蔓延到了学术界，触发了发展合作行业拥护者与反对者之间火热的争论。世界银行的前经济学家威廉·伊斯特利（William Easterly）——一位自称先驱的援助业批评者表示援助的问题在于其本身。在他2006年出版的《增长的古怪需求：经济学家们在热带的冒险与不幸》一书中，伊斯特利回顾了几十年来的发展项目并总结道，援助行业已经使国民生产总值较低的国家产生了寄生虫情结，并且这一现象由于援助者的“仁慈”再度恶化，援助资源统统落入了腐败政府的囊中。作为自由市场坚定的信奉者，伊斯特利坚持认为，援助行为本身的性质忽略了人们对于经济刺激因素的反应，由于源源不断涌入的发展资金，政府便不会认真思考如何促进系统的改革，市场化被延迟、真正的竞争（政治和经济两种层面上的）被扼杀。对于伊斯特利，援助只是一个“诱饵”而绝非强迫对象国家支持长期、深层次经济改革的“棍棒”，而且，取消对象国债务的风潮已经导致了道德危机。伊斯特利认为，援助行业是建立在糟糕的经济考量之上的，并且人们没有认识到，只有释放真正的市场动力、奖励良好表现以及鼓励竞争，才能够实现真正的发展成果。2007年，伊斯特利在

《白人的负担：为什么西方的援助导致了如此多的损失且鲜有成效》一书中再次提出了他的批评。

高盛前经济学家、畅销作家戴妙玉（Dambisa Moyo）提出了一个相似的观点：“因为鼓励腐败、滋养依赖性、助长通货膨胀、制造债务负担以及剥夺非洲人民的公民权利（举例来说），基于援助的战略造成的损失远多于益处。”她认为马歇尔计划一类的干预措施在欧洲的“重建中发挥了重要作用”，然而，她评论道，“这样的干预措施与今天在全球实施的援助之间关键（却常常被忽略）的差别在于，前者是短期、高效、有限的，而后者却是看上去永无止境的承诺。”戴妙玉认可伊斯特利的观点，坚定地认为此类无限期的系统没有向政府提供“任何刺激因素使其寻求其他更好的为发展融资的方式”。与她看法相同的还有英国经济学家保罗·科利尔（Paul Collier），他指出援助的收益在逐渐下降，因此（根据0.7%的目标显示）越来越多投向对象国的资金也难以增值。科利尔引用的一系列例子显示，在全部投向对象国的资金中，只有一小部分（约为1%~2%）实际上达到了预设的目标。

更多的乐观主义评论家则驳斥对援助如此消极的观点。2005年，哥伦比亚大学地球研究所主任杰弗里·萨克斯（Jeffrey Sachs）对发展援助发表了著名的“维护”言论，在《贫困的终结：我们这个时代的经济可能性》一书中，他论述道：“援助是非常重要的，如果利用得当，则能够实现长期可持续的成果。”萨克斯陷入了一场与伊斯特利和戴妙玉的私人口角，他谴责他们“兜售简单的自由市场与自救主义的混合配方”、都是援助行业的既得利益者（伊斯特利曾供职于世界银行，而戴妙玉也得益于为非裔年轻人提供的奖学金才能够在哈佛学习——该奖学金由援助行业提供），并且，他强调：“现实生活中，我们总会在某些时刻、不计其数的方面需要帮助，更重要的是我们应该考虑到，对于穷困的人来说，他们一旦失去了帮助，面临的则是生与死的后果。”

尽管援助效力争论的双方之间存在差异与既得利益（例如伊斯特利供职于世界银行期间曾涉足加纳的结构性调整项目，而萨克斯也曾为发展组织、大型慈善机构与国际慈善组织提供咨询服务），但他们都在同一点上达成了共识：援助行业需要基于数据的精密方法，因为数据能够提供有意义的总结并辨别有效方法。对于大多数争论者而言，这个行业只有通过精密的测量方法来优化资金分配，才能够实现长期发展成果。如果不厘清这一点，那么批评和褒奖则无法摆脱道听途



说与个人偏好。

## 对于证据的需求

许多援助行业的批评者指出了一些方法论方面的解决途径，包括应用实验工具评估干预措施的效力。在科学知识中，实验处于中心地位已经不是一个新鲜的话题。现代科学的奠基者，意大利人伽利略·伽利莱（Galileo Galilei）支持了实验方法的使用，并且将调查的科学方法重新定义为归纳探究，而非推理证明。1849年，物理学家迈克尔·法拉第（Michael Faraday）写道：“世间万物但凡遵循自然法则，那么即便它们再美好也都是实在的，实验能够最好地测试对自然法则的遵循。”查尔斯·达尔文（Charles Darwin）也强调了实验科技在农业与生物领域中的重要性。在生物医学领域也是一样，实验方法已经成为评估药物、方案与治疗方法效力的常用手段。该领域内最常见的试验方法是随机对照实验（RCT），这是一种通过在干预组内随机挑选参与者以避免潜在的因果关系的评估方法，干预组包含接受“治疗”的个体，而作为对照，所谓的控制组则不受干预。由于捐助者及其顾问对于使用多种经济手段进行“因果”推理很有兴趣，所以随机实验在发展领域以及更常见的经济、社会研究中应用已经如火如荼。（本书分享更多[索搜@雅书](#)）

尽管如今一大批随机对照实验专家在国际发展组织进行实验操作，但最出色的经济学家还是在麻省理工学院。2003年，贾米尔贫困行动实验室（J-PAL）在这里落成，该实验室“汇集了全球抗击贫困行动中使用随机评估来回答批判性政策问题的调查者们”，实验室的领导人是阿巴吉特·维纳雅克·班纳吉（Abhijit Vinayak Banerjee）与埃丝特·迪弗洛（Esther Duflo）——两位终身致力于随机实验与发展影响评估相结合的发展经济学家。他们的著作《贫穷经济学》获得了《金融时报》暨高盛年度商业图书奖，成为被引用次数最高的随机控制实验类文献。该书概括了15年来从全球实施的随机控制实验中产生的证据，强调了发展理论的必要性，这些理论必须“使我们理解穷人有能力实现什么，以及他们在哪些方面需要我们的支持”。诚然，贾米尔贫困行动实验室的调查者们认为“发展政策通常是基于潮流制定的，而随机评估能够使它的制定基础变为实证”。

带着“用硬数据抗击全球贫困”的目标，他们鼓励系统地应用随机控制实验以“产生对该领域内项目或政策真实影响的客观评估”，这样，便能够打破以往的悖论——发展项目每天都在消耗庞大的资源，而我们“仍然无法找到有效抗击贫困

的方法”。对于贾米尔贫困行动实验室的调查者们，贫困是不理性行为、不理想经济决策的后果，《贫穷经济学》阐述了“日常生活费用低于1美元的生活如何迫使穷人做出值得怀疑的决定——助长而非抗击贫困的决定”，在这方面，调查者主要负责研究贫困人口行为、他们犯的错误以及怎样的援助能够矫正这些错误。

与贾米尔贫困行动实验室密切合作的是智库“扶贫创新行动”（IPA），该智库领先地系统使用随机评估以“确定不同项目的实际影响与成本效益”。不同的是，贾米尔贫困行动实验室更具学术性质，而“扶贫创新行动”在科技援助方面强调实践方法，并且旨在将经验“传递给全球的政策制定者、实践者、投资人与捐赠者”。他们所说的也很大胆，在报告中，他们表示要“资助有效方案”、发掘“高成本效益”解决办法以及进行“精密的评估”。有趣的是，他们指出了一个完美的例子——微型贷款，“人们在发现这个概念奏效的证据前，它就已经激发了巨大的热情与支持力度”，即使这个结论是“基于评价高息债务有用性的双重标准预测得的”。他们基于实证，科学地审视非传统型方案，以确认这些方案是否证实我们预想的概念与期望。

“扶贫创新行动”提出的影响证明倡议具有更高的利益，这个倡议包含了他们遴选出的能够产生预设成果的发展方法，只有当某种干预措施“在至少一种情况下”被测试过、由“科学的精密方法（随机控制实验）”评估过，以及通过“成本效益测试”，它才能够入选这一倡议，并且，一个遴选委员会负责这些措施的科学审查和遴选过程：被指定为“影响获得证明”的想法将被直接进行推广，无须任何附加的操作调查，而被指定为“有待观察”的想法则需要满足附加标准才能够进行推广。这一倡议甚至还有一个专门基金会，该基金会公布反响良好的援助做法，并鼓励投资者们依照这些做法选择捐助目标。影响证明倡议的资助者迪安·卡兰

（Dean Karlan）在他最近出版的《不止于美好愿景：新经济是如何解决全球问题的》一书中回顾了援助成本效益争论中的主要阵营以及他们关于发展援助明显无法调和的批评与支持：“我有预感，到最后即使萨克斯和伊斯特利也会在以下观点上达成一致：援助有时奏效，有时失灵。”对于卡兰，关键的问题是：哪种援助才能奏效？其实答案并不在极端情况下，而是在一般情境下就能够找到——针对某一个困难、提出一个可能的解决方案，然后通过测量来检验它是否奏效：这才是消除贫困上“真实的、可测量的并且有意义的成果”。

尽管影响证明倡议的方法很大胆，但他们证明影响的都是传统领域。尽管谈

到自己的想法时他们胸有成竹，但是这些想法本身听起来并不怎么具有革新性，似乎都是合理的建议，但这些想法能够解决世界上最古老的社会不公平现象吗？

卡兰和他的同事认同影响证明倡议结合了传统与行为经济学的方法，他们表示，穷人也像所有人一样，“会做出使自己更穷、病得更厉害以及更不幸福的错误决定”，接着他们又开玩笑似地补充道：“如果他们不是这样的话，那么他们早就在向我们这些剩下的人出售自救课程了。”影响证明倡议同贾米尔贫困行动实验室一起，使用硬数据与科学推理来矫正这些错误。用农业来举个例子，如果化肥和“改良”品种（即转基因品种）能够提升作物产量，“不应该有更多的农民在这上面投资吗？”影响证明倡议的调查者在他们的网站上问道。回答是肯定的，但前提是在正确的时间释放正确的刺激因素。这些经济学家凭借心理学与经济学中的拖延症模式“了解到，许多人相比未来的消费更重视当下，他们倾向于拖延有利可赢的投资——即便他们确定自己有投资意愿”。由于农民的拖延症，他们面临着错过“化肥或其他农业投资在一季中可获得利润的时间”的风险。这种内在不稳定性可以帮助我们理解投资的真正价值，并且也是“为什么商贩要挨家挨户、在一季的丰收后（这个时候农民有更多的钱）马上推销化肥，推销手段诸如限时折扣与免费配送提前付款的化肥等，这样就抵消了拖延症一类的行为倾向”。影响证明倡议采用这一推理过程，鼓励凭证系统的使用与发放智能补贴以“补贴个体下一阶段的预算，像即将来临的种植季补贴，或者为学生发放的返校季补贴”。

卡兰和他同事的观点与最传统的成本效益分析支持者相比并没有很不同。根据以上方法，唯一能够让个体行为变得理性的便是投资凭证、限时补贴与其他形式的资金劝诱。有趣的是，这些说法却与另一位硬数据支持者——咨询顾问布里奇斯潘的研究结果背道而驰，该研究显示的最优实践来自非营利性组织“直接捐赠”（Give Directly）。这个组织获得了2012年“全球影响力奖”，该奖项是谷歌支持企业非营利性行为的一个项目。通过随机对照实验，“直接捐赠”发现帮助人们的最好方式就是直接给钱。“直接捐赠”组织的“严格的方法”使他们至今都不赞同影响证明倡议与贾米尔贫困行动实验室的做法，他们坚持认为，无附加条件的直接捐赠能够改善健康与未来的经济收益，因为穷人愿意在方方面面投资，从给挨饿的孩子的食物到长期资产包括土地、牲畜与房屋。他们认为，穷人十分理智，他们知道自己需要什么以及如何得到这些东西，他们不需要发展顾问来帮助自己理性地行动。到底谁才是对的呢？同样的数据难道可以产生相反的结果吗？

有趣的是，尽管影响证明倡议拒绝采纳简单的经济发展方法，但他们的解决办法与第三章中提到的哥本哈根共识中心所提出的解决办法并没有太大不同。哥本哈根共识中心汇集了世界上首屈一指的用“科技办法”解决全球问题的科研人才。2008年，由诺贝尔经济学奖得主组成的哥本哈根共识中心“梦之队”发现，微量营养素干预——用于增加营养摄入的营养强化和补充手段，是改善发展中国家营养不良最有效的投资项目，该项目能够“通过低廉的价格获得巨大的利润”，也与影响证明倡议的某些方案相契合。2012年，哥本哈根共识中心的调查人员表示，每年只需要不到7亿美元“就可以解决全球饥饿问题”。他们的解决方法与影响证明倡议的相似，包括：“营养干预包、增加全球食品产量，以及通过改善沟通与强化化肥市场的竞争来优化其市场功能”。当然他们也注意到“增加全球食品产量是一个奇怪的想法，因为全球范围内食品总是供大于求”，然而他们坚持，“降低食品价格让更多的人负担得起”，以及“为气候变化的恶劣后果做好缓冲工作”。毫不意外地，他们的结论与食品行业的关键企业利益相符合：通过广泛育种提高产量，增强作物对干旱、高温与盐的耐受力，辨别并传播最优良的作物品种以及确保化肥的最佳使用。他们还建议通过短信息向农民发送市场信息（按月计费）以及让农民更方便地得到化肥，比如上门推销。哥本哈根共识中心的经济学家们表示，这些“创新”能够带来高达成本“8.35倍”的收益。

化肥、转基因种子、短信提醒、上门推销：多亏了硬数据的分析方法，哥本哈根共识中心的宏观经济分析才得出了与影响证明倡议在微观个体行为层面所总结的相似的结论。当然，哥本哈根共识中心的调查人员也认可：由于“少数国家掌握了绝大多数氮、磷肥与钾肥的生产”以及“四家龙头企业掌握着整个国家一半以上的产能”，所以“旨在鼓励可持续化肥使用的政策产生了混合后果”。但是为了解决这一问题，他们需要警告政策制定者避免“强制性打破这一垄断产业”，因为这会对产业造成“严重破坏”，并且会“导致规模经济的消失”，除此之外，还会阻碍行业监管，因为强制打破会造成零收益的寻租行为。相反，调查人员建议对将“用于新产能建设的公共投资转向私人部门”，目标是在亚洲与非洲建立两个大型企业集团，总耗资预计约13亿美元，这两所亚洲与非洲的“孟山都公司”将为当地提供化肥与转基因种子，预计净利润约125亿美元。

影响证明倡议与哥本哈根共识中心忙于计算成本效益比率，却未提到许多调查在过去几十年内阐述的一点：市场驱动的对本地农业的破坏是导致当今世界大部分地区“无法保障食品安全”的原因。

印度就是最好的例子，过去印度农民只能依照自然时节种植作物，好时节带来好收成，当季的粮食和种子盈余就用来减轻恶劣年份造成的影响，男男女女依照着自然规律。这也意味着食品生产只停留在满足生存需求上，农民也不为自己致富，因为他们是规定的“穷人”，但是除非有大型自然灾害降临，一般情况下村庄里都有足够的粮食来维持生计。后来，在20世纪80年代早期，印度政府开始听取援助专家、发展经济学家与食品公司的建议，决定利用新兴科技扶贫的潜力，于是，印度引进了转基因（GM）种子，将农业从维持生计的活动向大规模生产转型。自那以后，商贩便挨家挨户推销“改良”种，第一批出售或许是像影响证明倡议专家建议的那样，每逢季末——在农民较为富余的时候打折出售。

大规模的广告宣传让农民相信，他们可以花很少的力气生产出更多的粮食，因此，接下来的几十年，整个印度都将天然的作物种子换成了转基因种子，这些转基因种子的专利所有者和出售商都是全球最大的跨国公司之一——孟山都。印度农民寄希望于这些新种子来抵抗多种寄生虫，包括让人堪忧的水腊虫。因为转基因种子无法自然获得而只能在商品市场上买到，所以农民不得不贷款来提前准备好购买种子所需的资金。而在一个银行都厌恶穷人的社会中，许多农民只能接受高利贷者的条件，这往往要求他们转让自己的土地所有权以防欠款不还。这个时候局面便开始急转直下了，由于转基因种子的效果并不理想，其粮食产量相比自然（零成本）手段带来的粮食产量也没有高出多少。于是，农民的处境便更糟糕了，他们意识到自己的收入无法偿还债务、养活家人，同时也不够买下一年耕种的新种。从短期看，成本始终高于利润。意料之内的是，印度这样一个半数以上人口依靠耕种生活的农业国，因其激增的农民自杀率而变得臭名昭著。大多数对印度农民的自杀人数的估计是每年17000人。根据印度国家犯罪记录局的调查，1997至2008年间印度自杀总人数高达199132人。然而这些数据恐怕无法描述悲剧的规模，因为在一开始，只有拥有明确土地所有权的个体才能被“算作”农民，这也就默认地排除了女性农民和租赁土地的农民，而且“农民”这一概念还包含职业农民与偶尔进行农事活动的个体，因此，“这些数据折磨着我们，它们低估了真实的自杀人数，却高估了真实的‘农民人数’”。

这些农民的自杀原因并不是天气灾害、更恶劣的虫害或气候变化的影响，他们的自杀是被沉重债务夺去尊严后的一种极端反应。雪上加霜的是，他们还失去了获得社会地位与收入的唯一来源：土地。并且，农民人数的减少并不只是由于自杀，他们当中的许多人抛弃了村庄而迁移向城市寻找立足之处，沦为印度城区

庞大贫民窟的一员。最后，农民自杀、土地剥削与种子商品化都成了公共资产私有化过程的组成部分，而这一过程却被夸大为印度从落后国家转型成为世界引擎的前提条件。

2003年，英国最大的专业工程协会机械工程师学会的报告显示，全球三分之一的食物（约20亿吨）“从未上过人们的餐桌”，这些食物由于“市场与消费浪费”被直接浪费掉了。这一现象在工业化程度较低的国家尤为广泛，在这些国家，当地社区被土地所有权制度与农业商业化夺去了土地和粮食。而在所谓的发达国家，市场导向下的食品浪费非常普遍，大型超市惯例上不接受任何“不符合具体市场标准的——如体积与外形不达标的食用果蔬”。在英国，由于这样的现象，高达30%的食物都被浪费了，而在全球范围内，最大的零售商“以这样的方式每年浪费高达160万吨的食物”。而且，市场策略与营销手段鼓励消费者过度消费，于是便导致了家庭浪费：在发达国家每个消费者会将其购买的30%~50%的商品丢掉。机械工程师学会总结道，控制与减少浪费“远超出了个体农民、商贩与消费者的能力范围，因为那是由市场规律决定的”。

于是，不可避免的问题便是：随机实验能够在多大程度上显示以上事实？实验方法与成本效益分析又能够多大程度上暴露这一导致贫穷的深刻动态？如果新的经济革命真的能够解决全球问题，那么在一个由权力动态和失败的制度所主导的社会中，问题的解决不应该从调查对象的重新布局开始吗？恰恰相反，强调实证正导致援助行业更加依赖技术治国手段与短期收益。《可怜的数据》一书的作者莫滕·杰尔文（Morten Jerven）指出：“在实验环境下研究这些问题也许会误导学者与政策制定者们，说不定，最重要的其实是政治经济与实验之间的差别而非相似之处。”该书阐述了数据如何误导发展中国家的政策制定，以及随机实验无法暴露出社会生活当中的“政治层面”。

## 慈善资本主义的崛起

对硬数据的强调以及它在效率与效力（这是一个很大程度上从商业术语中引用的概念）方面的滑坡使援助行业呈现出一种新趋势，这种新趋势可以被恰当地称为“慈善资本主义”，这个概念在2006年由美国《经济学人》编辑马修·毕索普（Matthew Bishop）提出，后来成为2008年一本畅销书的名字：《慈善资本主义：富人在如何拯救世界》。讽刺的是，该书上市后的几个月，富人们并没有拯救世界，而是陷入了深渊：雷曼兄弟银行破产、全球资本金融衰败。慈善资本主义的概念很简单：为了抗击贫困，我们必须让市场经济的佼佼者及其规律替代援助行业。由于几十年来的传统援助政策并没有产生实实在在的成果，全球市场及其商业龙头们又积累了史无前例的资产，单单是这一鲜明的对比——正如慈善资本主义教条所描述的——就足以让商人们理所应当地解决全球最令人担忧的问题。怎么解决呢？当然是重新发明手段来评估如何做有用——就像随机实验革命的原理一样。

正如书中的概要所说，“（作者）对全球最富有的人进行了访谈……描述了一群天南海北的施予者，正满怀热忱地准备改变世界”。这群也许是世界上最富有的精英们也被称为“好人俱乐部”，其中的慈善资本家们包括比尔·盖茨、沃伦·巴菲特、理查德·布兰森以及美国前总统比尔·克林顿。纽约城市大学的社会学家罗宾·罗杰斯（Robin Rogers）的观点是，“如果想要弄懂慈善资本主义，那么你必须先知道三个M：金钱（money）、市场（market）与测量（measurement）”：

第一个M，金钱（money），代表着财富，尤其是那些超级富豪们，应该用他们的财富创造更多公共利益……第二个M，市场（market），是说市场的力量应该将有效的社会项目与低效的社会项目区分开来。第三个M，测量（measurement）指我们应该基于数据有目标地、理性地利用资源，以辨别并推广成功的社会项目。

慈善资本主义发展观的一个核心因素就是解决问题：他们认为贫穷是一个独立的问题——也就是说，与复杂的社会经济结构无关——因此他们寻找创新方法来“解决这个问题”。慈善资本家认为在他们采用的商业框架中，自己比一般的企业家们更加出色，因此他们坚信能够找到更快、更好以及更有效率的方法来解决



根深蒂固的社会问题，比如，盖茨基金会（Bill and Melinda Gates Foundation）已经成为卫生保健及相关领域如营养、水、公共卫生和农业发展的龙头发展组织。盖茨基金会凭借超过360亿美元规模的资产信托基金（以及每年向保健部门投入的相当于世界卫生组织全年预算的捐款数量），已经成为世界上最大、最具影响力的慈善组织。在“致基金会的一封信”中，盖茨夫妇解释道，他们的工作重心在于研发“能够帮发展中国家农民收获更多粮食、挣更多收入的新科技、预防并治疗致命疾病的新方法、帮助学生和老师的教学手段”。比如说，在农业部门，他们强调“改良”种、优质土壤和牲畜非常重要，除此之外，他们还希望帮助农民“锤炼商业管理技巧、获得更强大的购买力与市场影响力，以及提高他们的作物与资源管理技巧”。2003年1月，该基金会发起了名为“全球健康大挑战”的倡议，旨在推动科学研究“解决关键科学与科技问题以实现发展中国家疾病治疗的进步”。这些大挑战包括实现一系列科技目标，从产后就能立即注射的单剂疫苗到量化评估人口健康状况的科技。的确，该基金会一直为获得实证基础上的成果不懈努力。在2013年1月，比尔·盖茨在《华尔街日报》一篇名为《我解决世界上最大问题的计划》的专栏文章中强调，测量方法对于改善人类生活十分重要：“如果你有一个清晰的目标，还有一项能够驱动你向目标前进的测量方法，那么你就能够取得出色的进步。”比尔·盖茨回顾了他的慈善事业所活跃的各个领域，他认为在过去十年中，由于新型测量方法与商业手段，世界上最贫困人群的生活已经出现了迅速的改善，并且他总结道：“多亏了测量方法，进步不再是什么稀罕、难以捉摸的事情。事实上，我们可以让它变得更常见。”

“保守”的福特基金会（Ford Foundation）前主任，或许是最直言不讳的援助行业新趋势批评者迈克·爱德华兹（Mike Edwards）表示，“好人俱乐部”的领导人们将自己包装成冒险的慈善资本家，强调他们的社会责任基于市场方法，持续谈到“新型的”“参与性的”“战略性的”“有效的”或者“有影响力的”慈善。“但这些话术并没有一个实在的概念那样有用——它们听起来太普适了——除非真的有哪个基金会追求的就是冷漠和低效”。事实上，冒险慈善这个概念本身就很具有迷惑性。“冒险”一词究其词源表示极大的不确定性，以及在风险前仍愿继续其事业的意志。相比之下，这些新慈善家们的花言巧语渗透着“高效”“成果基础”与“数据驱动”等概念，并且他们将成本效益分析作为工具来决定在哪个领域以及如何实施干预、为什么要坚决避免风险以及尽量规避更具“政治性”的领域。“拼搏集团”是一家华盛顿的慈善咨询机构，其创始人及董事长亚当·沃尔德曼（Adam Waldman）认为，新慈善诞生的标志是“基于成果的企业框架、杠杆、个人参与与浮躁”。世

界妇女基金会前董事长卡维塔·拉姆达斯（Kavita Ramdas）认为，这些新慈善背后的主观因素正是曾经造就了那些对冲基金经理、金融巨鳄、信息通信科技企业家与软件发明者的“问题导向”心理。他们进行国际发展的方法能够“产生可测量的、便捷的解决方案”，这样的能力赫然体现于他们的专业资历中，也正是那些“大型慈善机构”所需要的：“经理、顾问、工程师、商人、前行业领袖或游说家”。他们只强调效率与部门科技能力，但是“在这个测量驱动、追求效率以及强调科技的社会变革方法中，与人、人类和社会复杂性、文化和传统现实息息相关的发展似乎显得不那么重要”。摩根士丹利私人财富管理公司的慈善管理总经理、董事长梅拉妮·施诺尔·别贡（Melanie Schnoll Begun）认同这一观点，她认为“我们所打交道的这个时代无止境地渴求速效，他们一心只想看到成果……这公平吗？不，这对吗？不。发展组织需要后退一步，告诉慈善家们测量结果有多困难”。

因为可持续发展是一个旨在设计、巩固文化纽带、经济政治制度的长期过程，而不是一系列成功的援助项目，所以慈善资本家们对数据趋之若鹜的浮躁会严重地削弱全球国家实现可持续公平的成果的能力。在许多方面，目前对于援助效力的争论和基于成果的方法都会造成影响生产力的趋势，因为这一争论的重心已经从我们需要什么转移到了我们可以测量什么。这一趋势同样会影响官方发展机构的运作和工作优先次序，这些官方发展机构与我之前提到的欧洲援助组织相似，承受着来自国内选民们的压力，他们要创造出“可见的、可计量的以及可直接分配至捐助活动中去的援助成果——即使这么做会削弱援助的发展效力”。这就是为什么“测量有效方法”哲学的支持者们至今还认为，发展资金应该开放地向能够通过基于数据实验方法的“硬”实证测试的干预措施流动。由于大多数发展项目的受益者都难以发出自己的声音，因此，相比倾听这些人的需要，生产数据以取悦纳税人与投资者变得更加重要。基于成果的发展工作越来越重要，以至于数据模型已经成为资金分配的决定因素。拿随机控制实验举例，并不是所有发展干预措施都能够被随机化，涉及经济改革、制度发展、社区营造等方面的宏观项目几乎不可能用标准化模型来评估。一般情况下，一项干预措施越是庞大复杂，那么设计出一个评估它的模型就越困难。在这方面，标准化的支持者猛烈抨击了发展部门中的新趋势，比如说国家预算支持。但是，人们需要知道，国家预算支持（通过调整对象国政府工作先后次序来引入援助）本来是发展行业内部缺乏协调、出现分裂时的专门响应措施，而且，发展部门对于伙伴关系与对象国当地所有权的重视，也对对象国排列的与捐助国支持的国家优先事务之间的统一提出了要求，

因此，相比个体项目，国家预算正获得越来越多的资助。但是，像班纳吉与迪弗洛这样支持实证基础上的方法的人却将国家预算支持视作“灾难”，因为它把各种各样的资源集中到一起，使得援助与预期成果之间的因果关系无法厘清。这样的批评也对捐助者产生了影响，尤其是新慈善组织（也包括世界银行），越来越不愿意资助那些无法用“硬”数据基础上的实验测试评估的项目。“在最后”，一些分析者警告道，“这样的方法只能够提出问题，却无法通过分析知识缺口寻找答案。”而且，硬数据由于强调易观察的变量，往往会忽略更庞大、更有影响力的动态因素，标准化模型的应用很有可能用“显而易见的答案回答错误的问题”。而真正重要的东西，比如文化学习、社会信任与制度发展，当人们发现它们无法像给儿童疫苗、由小型信贷项目资助的新型企业、分配给当地村庄的防疟疾蚊帐数量或小学考试中学生的分数一样被统计时，都变成了次要因素（甚至被当作外部效应）。

很明显，发展行业需要在短期内的切实成果与长期进行基层培育的不确定性和风险之间进行权衡（在追求援助效力的时代，慈善资本家们饱受议论的“浮躁”已经成为发展行业的魔咒）。营利性的援助机构，无论是迫于来自选民（纳税人）的压力还是来自改善自己公共形象的压力（正如大多数新的大型慈善机构所面临的），都更有可能选择短期项目与可证实的成果。要求承担责任——这一公共辩论中非常受欢迎的原则，却由于带来成果驱动的短期主义而吊诡地造成了消极的影响。在数据驱动的技术治国论视角来看，发展是现有经济、政治与社会框架内技术操作的成果，因此，更深层次、更激进的社会变革的可能性才会被忽略。

有人可能会非常疑惑，在有史以来最大规模的救市和公开收购私人企业面前，商业技术优先于政府与慈善机构这样的观点怎么还有（越来越多的）人认同？的确，有人可能会厚颜无耻地指出，慈善资本主义宣言出现得“不是时候”，它出现在2008年，几周后便发生了华尔街危机。但是，危机持续恶化，传统援助行业缺乏资源，这对于商业与市场方法进入援助行业是更好的时机。在全球经济衰败的背景下，许多发展组织、非政府组织与非营利性组织都在努力维持运营，还有，当然，在这个公共预算缩减、政府难以满足公民要求（更不要说关心海外发展目标了）的世界上，有越来越大的空间让全球的亿万富豪们采取私人行动，正如《慈善资本主义》的作者所说：

2008年金融危机带来的财政影响说明，公共预算与政府的野心需要缩减至少一个世代……慈善资本主义革命有着非凡的意义。政府正在削减在社会事业上的花销，这也许会为世界上的社会变革提供前所未有的动力。

对成果与数据的重视已经改变了全球慈善业运作的方式。慈善家们变得越来越像投资人，他们盯着一组数据，计算着发展项目的可行性，于是越来越多人对于极具吸引力的“价值”评估提出了需求。这一新兴市场需要新的基础设施：与股市性质相当的慈善机构、投资银行、调查所与管理顾问，这一基础设施不再基于原则、任务与价值判断而是真实的评级。在开支审查与公共财政紧缩风行的这个时代，三个M（金钱、市场与测量）涉及的范围不再局限于低收入国家的援助政策，而是凭借其强大的力量进入更富有的国家。

## 数据的罢工：社会领域影响评估的政治

1957年，社会学家唐纳德·坎贝尔（Donald Campbell）出版了书籍《社会环境中与实验有效性相关的因素》，该书后来对社会科学产生了持久的影响。在书中，坎贝尔介绍了许多如今已成为社会调查方法论中常见术语的概念，包括内部/外部效度以及（类）实验设计。他偏好使用随机化技术，但是作为一个现实的人，他很清楚，由于资源与现实情况的限制，类实验需要满足绝大多数情况下的需求。他在“实验社会”方面的著作成为指导世界“寻找可普遍推广的有效改革”的蓝图。他强调“硬”评估技术，因为他坚持认为科学研究总是能够发现因果之间的清晰关联。在学术界，坎贝尔已经成为社会领域实验方法与实证主义验证的标志人物。2000年，以他命名的“坎贝尔协作会”成立，该组织专门对教育、犯罪与公正以及社会福利等社会干预措施效果开展系统审查。正如贾米尔贫困行动实验室在国外援助项目中提倡随机控制实验，“坎贝尔协作会”也是美国提倡以这些实验手段评估社会工作最直言不讳的支持者。

影响评估不仅仅是社会科学方法论争论中必不可少的部分，还是政府最开始系统理解公共服务项目效果的努力。举个例子，1964年美国政府发起了“向贫困宣战”协议，该协议后来促成了经济机会办公室的成立，建立了一个专门的调查评估团队。1968年，该办公室在新泽西进行了一项测试负所得税实际可行性的田野实验，该实验后成为历史上第一个使用随机控制方法的大型社会科学实验。同样的，在环境影响评估领域，各类成本效益分析与社会影响评估被引入《美国国家环境政策法》中，美国社会学协会委员会对此表示了支持；1973年，美国内务部在起草《环境影响报告书》时首次使用了“社会影响评估”一词；1974年，福特基金会与其他一些联邦机构建立了“人力评估研究公司（MDRC）”，该公司是一个通过引领随机化技术以影响全国立法、项目设计以及实践操作的非营利性调查组织。“人力评估研究公司”在其官网上声明，他们的工作是改善“情绪、意识形态往往左右辩论的领域”，并且致力于“客观公正地证明高成本效益解决方案，以进行复制推广”。1980年，社会影响评估开始成为世界银行发展工作的一部分，并且在20世纪90年代，美国政府成立了“社会影响评估指导方针和原则国际委员会”，该委员会专为联邦机构的评估方法开发标准与要求。

自那以后，基金会、非营利组织、风险慈善家与社会投资者们都开始采用评

估方法，用一系列模型、方法与公式来进行核心影响评估实验。1997年由美国投行贝尔·斯登（Bear Stearns）的前金融家乔治·罗伯茨（George Roberts）建立于旧金山的“罗伯茨企业发展基金会”（Roberts Enterprise）在这方面扮演着突出的角色，罗伯茨以完成了全球金融史上最大规模的杠杆收购而闻名——他于1989年以250亿美元的价格收购了雷诺兹·纳贝斯克（RJR Nabisco）烟草公司。杠杆收购指的是凭借“友好”银行所发行债券（在某些情况下这些债券都不是投资级别的，而是所谓的“垃圾债券”）获得的大量贷款（通常以90%的贷款对应10%的净抵押资产，但是也有过100%贷款的情况）而完成的收购。由于目标公司当下的成功（或是预期内的前景）被虎视眈眈的收购方视作抵押品，杠杆收购往往是一种残忍的、掠夺性的策略，尤其会被用来收购或者摧毁竞争对手，因此一个公司越是成功，就越有可能遭到“攻击”。美国证券交易委员会（SEC）前首席经济学家G.A. 贾雷尔（G.A.Jarrell）表示，杠杆收购的弱点在于“价值缺口”，即一个公司当下价值与其股票的更高预期价值之间的差距，这一差距往往导致目标公司在成功被收购的情况下，或是在防止自己被收购的情况下承担超额债务。杠杆收购的批评者们包括美国联邦储备委员会保罗·沃尔克（Paul Volcker）与美国证券交易委员会前主席约翰·沙德（John Shad），他们保守地认为这种债务支撑的收购会为公司股东、债券持有人、员工、顾客、供应商、当地社区以及纳税人带来风险，因此增加金融危机、资产损害与失业的可能性。纳贝斯克事件标志着恶意金融收购史的开端，美国《时代》杂志在1988年为此奉献了一整个版面（恰当地为该文章起名为《贪婪游戏》），电影行业也于1993年拍摄了影片《门口的野蛮人》，该影片抨击了逐渐逼近的金融风暴与受其牵连的华尔街投资银行。

在这一背景下，意料之中地，是罗伯茨的非营利性组织“罗伯茨企业发展基金会”引领了利用金融模型评估社会项目的做法，尤其向慈善家与非营利性组织引入了“社会投资回报”的概念。“社会投资回报”遵循成本效益分析原则，测量一项干预措施（或是整个组织的活动）产生的社会效益价值，然后与需要实现这些效益所需的投资规模进行比较，来评价这一干预措施的好坏。“罗伯茨企业发展基金会”将社会（以及环境）回报用货币价值来衡量，旨在在社会领域使用“硬”经济数据来展示价值创造。

“社会投资回报”一般可以这样形容：

## 当前净效益价值

### 当前净投资价值

这一结果是货币化后价值的简单比率，比如，4:1的“社会投资回报”比率代表1美元投资带来4美元社会收益。这个方法相当简单，并且对全世界投资者有强大的说服力。如今，学者、评估者、顾问以及社会工作者正在进行无数种“社会投资回报”分析，他们当中的一些人具有明显的“评估”天分——也就是说，他们负责对过去的干预措施进行评估，相比之下，另一些人更加有前瞻性，即他们负责预测新干预措施可能的社会价值，并且主要负责那些期待通过投资获得惊人社会改善的慈善家与投资者们。有些情况下，“社会投资回报”也被用作可行性研究（在这一情况下“社会投资回报”也被称作“盈亏平衡”分析），来说明特定干预措施的社会收益超过其成本需要达到什么条件。

根据之前章节所述，“净现值”是一个饱受争议的概念，它是通过未来收益的折现率计算得出的。它假定“人们总希望在今天而非明天获得收入，因为未来存在风险（比如收入无人支付）或是因为未来存在机会成本（比如将财富投资在其他方面有可能产生收益）”。《2012年“社会投资回报”指南》显示这一方法“通过贬损未来以鼓励短期主义”并且“违背了人们其实珍视自己及子孙的未来这一道理”。然而折现率已经广泛应用于所有的“社会投资回报”分析中。比如，英国财政部《绿皮书》的建议基本比率是3.5%，这一比率为所有政策项目的考核和评估建立了框架。

在欧洲，英国的“新慈善资本”（NPC）已经成为“社会投资回报”领域公认的龙头组织，他们认为，思考“社会投资回报”的一种方式“就是了解利益相关者是愿意直接获得资金还是获得慈善组织提供的服务”：

比如，给一位领抚恤金的人100英镑，在短期内就能够让他的生活有显著的改善，他可以用这笔钱购买服务或食物，或享受这笔钱带给他的额外的经济安全感。但是，把这100英镑给一家为领取抚恤金者提供社会活动的慈善机构，也许可以帮这个人培养并维持与他人的长久友谊。许多接受慈善组织帮助的老年人都表示他们所接受救助的“社会价值”远超过100英镑。“社会投资回报”的目的就是量化这一价值——显示出这一价值多大程度上超出了其经济成本（如果它的确超出了成本）……通过用金融术语描述慈善成果来观察特定的慈善活动是否值

得我们为其付出的成本。

在“新慈善资本”的运作中，测量成本与收益至关重要。他们认为，尽管“我们都知道生态环境很重要”，“但是直到环境经济学家们成功对环境估值，并且将环境的价值与环保成本做过比较，政府才有动力抗击全球变暖。”“新慈善资本”将“社会投资回报”类比为环境治理成本效益分析令人费解，因为根据第三章与第四章，实际情况并非如此：直到经济学家决定使用成本效益分析来评估环保法规的优缺点，政府才在逐渐增加的社会运动与公民社会压力下引入这些法规，经济学家的分析也为市场基础上的环保机制如排污权交易开辟了道路，从那时起，一系列关键环保方面便难有进展。

“社会投资回报”分析一般遵循一系列预设步骤。调查者会确定目标社区（也就是利益相关者）并且明确投入与产出。我们发现，特定干预措施涉及许多传统因素比如员工薪资、志愿服务时长、租赁成本与其他形式的投资。在产出中，调查者会对目标人群数量、所提供的服务量以及其他与这一干预措施直接相关的因素进行评估（比如，如果调查者是在分析一项教学项目的话，他们就会评估学校的表现）。为了收集这些信息，他们需要向利益相关者发放问卷，并且采集额外的“硬”数据（比如卫生健康项目中目标人群的健康报告）。理想的情况下，“社会投资回报”分析应该能够追踪直接产出（即干预措施的直接成果）与“间接”成果之间的联系。直接产出与“间接”成果能够更广泛地描述该干预措施最后对目标社区的影响。在文化水平项目中，产出应该包括参与者的接受教育时长、招生人数以及考试成绩，而成果则包括学生的社会经济、文化与生理健康的整体福利。相比之下，成果更难进行评估，更不要说直接追踪受评估的干预项目了。那么当然，一切（投入、产出与成果）都必须被货币化，在这方面，“新慈善资本”建议采用“粗暴的”支付意愿研究，因为这些研究能够更好地形成“利益相关者的看法”，这也是“社会投资回报”的关键目标之一。但是我们并不清楚他们所说的“粗暴”是什么意思，因为根据第四章，支付意愿调查的相关问题至今仍然争议不断。

“社会投资回报”的支持者们都意识到，“直接评估不同利益相关者眼中干预措施成果的价值并不简单”，这也是为什么他们经常使用金融变量。该方法是基于服务成本（比如犯罪的减少所节约的警力或健康状况改善所节约的医院床位总成本）或市场价值（比如住房项目所避免的食宿成本）的价值评估。在英国，内阁办公厅与苏格兰政府始终致力于发展用于标准化“社会投资回报”分析估值过程以



及提升该过程信度的指标与金融变量数据库。“社会投资回报”的两项原则是“重视关键事物”与“只考虑物质资料”，他们偏好投入金融成本，“这样才能够显示成果价值”。然而，使用金融变量会不可避免地削弱“社会投资回报”评估反映利益相关者看法的能力，因此一些评估者偏好进行“社会投资回报”分析但是不给出干预成果的具体金融价值（这种情况下，专家们所讨论的是成本效用分析而非成本效益分析）。全球报告倡议组织提出了一系列强化测量非金融价值的检测与评估工具的建议，2010至2012年间我所供职的（如今我仍为其工作的）德国海德堡大学的社会投资中心研发了一系列使用“社会投资回报”框架进行价值货币化的方法。

显然，传统的“社会投资回报”评估方法会导致严重的问题。尽管测量一个住房项目的金融价值相比之下并不容易产生争议（即使有人认为“家”的价值高于该房屋的市场价格），但是用数据描述赋权、解放或人权教育等的价值是一件很难的事情，因为它们涉及许多定性的、微妙的非市场因素。不可避免地，我们发现这里与其他基于实证的工具比如随机实验的分析具有相同程度的目的性。而且，“社会投资回报”必须包含“本应该出现什么”的概念，但是这种反事实的实证明显不可能出现，并且，不论是怎样的概率或预估都会轻易导致计算误差。这些全部的方法论限制与产生不统一数据的风险都使得不同组织的“社会投资回报”比率无法进行比较，因此也限制了不同组织彼此学习、评估最优方法的能力。与此同时，测量投资收益又是一项极耗资源的工作，大多数非营利性组织与社会企业都趋向于将这类测量视为累赘而非竞争优势来源或是一项有益的活动。“社会投资回报”要求一个组织具有系统的、不受资源（比如人力、时间与资金）限制的数据采集机制。反过来，这又产生了一个无法避免的道德问题：社会工作应该更强调帮助人们还是更强调研究他们？

尽管存在种种明显的短板，“社会投资回报”还是越来越受欢迎，因为它对于乐于将金融市场的分析框架应用于慈善业的捐助者与投资者们来说十分具有吸引力。如此说来，高盛基金（Goldman Sachs Foundation）——高盛投资银行的资助性组织，在慈善业扮演着领导角色便不足为奇了。2003年，该基金会于其纽约总部召开了与资助者们讨论影响评估发展未来的会议，全体与会者（他们当中的大多数都是风险慈善家与投资银行家）一致认为传统的非营利性工作时代已经过去，社会事业的新动力将通过采用金融部门的最优做法及操作方法来实现，会议报告强调：

过去的几十年见证了从项目关联的资助向强调测量与成果的风险慈善投资的过渡……投资者们如今强调更高的透明度与责任，他们希望了解自己的资金对于世界的影响。

虽然人们需要学习才能了解“社会投资回报”，但很明显的是，很多“社会投资回报”支持者的根本目标是为根据传统金融市场制定的社会目标创造一个投资市场。“新慈善资本”创始人加文·戴维斯（Gavyn Davies）与彼得·惠勒（Peter Wheeler）是高盛的合伙人，他们在一次于伦敦该公司餐厅中进行的谈话中，产生了为慈善家们成立服务公司的想法。回想他们成立“新慈善资本”的决定，戴维斯说：

在20世纪90年代的金融市场，所有人都致力于在可能获得最高收益的领域进行投资，既然如此，为什么同样的事情不能出现在慈善业呢？我们发现，我们缺乏足够的通过脚踏实地、独立与高质量的方式产生的信息。

在欧洲，英国一直是“社会”市场革命的先锋，尽管“硬”测量热潮在新工党任期时就已经开始了，但保守党领导的戴维·卡梅伦政府进一步坚持采用帮助社会公司与私人投资的工具，因为这与其“大社会”的理念相契合。在这一领域，政府除了其他倡议还发放了“大社会资金”，一笔6亿英镑、用于支持建立全面的社会投资市场的资金。所谓的“社会影响债券”（SIBs）也产生了越来越多的利润，该债券是一种基于援助成果的缔约，在实现预期社会成果的情况下，向解决社会问题的项目投资的慈善投资者便能够获得金融收益。“社会影响债券”正在接受一系列项目，比如降低短期囚犯再犯罪率、减少接受惩戒服务儿童人数、开发更有效戒毒项目的测试。

德国菲纽斯（Phineo）公司遵循“新慈善资本”的方法，为非营利组织开发了一种评级方法，用于引导社会投资者并迎合他们的慈善计划。菲纽斯及其伙伴的工作与信用评级机构的分析相似，他们对非营利组织进行质量评估，并将最高等级（他们称为“影响力标签”）奖励给那些有可能产生最高投资收益的组织。社会投资领域的意见领袖，如曼哈顿研究所的霍华德·赫克（Howard Husock），始终致力于建立非营利“股市”。《数百万桶中的数十亿滴》的作者史蒂文·戈德堡

（Steven Goldberg）建议遵循风靡网络的市场预测领域与选股网站的模型，采用一种“影响指数”，人们通常通过这些模型预测球赛胜负、政治选举、股票价格甚至名人的死亡。戈德堡注意到这些形式的集体预测非常准确，他提倡使用同样的

方法评估慈善事业的价值。支持戈德堡想法的慈善资本大师马修·毕索普及其合著者迈克尔·格林（Michael Green）提出，“IMPEX（影响指数）的主要作用是利用群体智慧评估非营利性组织表现并将其排名，向市场提供大量信息，让捐赠者们知道他们的钱在哪些地方可以发挥最大的作用。”

## “行善”市场化

我们面临着一个明显的悖论。对援助效力与影响评估这两个本身崇高的目标的追求，最终使社会变革沦为了市场效率的附庸。而透明性与责任是两个无论怎么看都具有价值的原则，但是无论在国际层面还是国内层面上，用于提高透明性与责任的工具方法都沦为了市场化慈善业策略中的“特洛伊木马”。也许这已经并不新鲜了，我个人始终对慈善的本质保持怀疑态度，作为小说家约翰·斯坦贝克（John Steinbeck）的仰慕者，我同意他所说的：“或许在全部的伪道德当中，给予是最被高估的一个。”

给予塑造了施予者的自我心理，让他们觉得比接受者更优越、更高人一等以及更强大。几乎在任何时候，给予都是一种自私的愉悦感，并且在很多情况下都是一种显而易见的、具有破坏性的以及邪恶的东西。一个人只需要记住，某些贪婪的金融家花费他们一生三分之二的时间从社会中攫取财富，再花费剩下三分之一的时间将这些财富还回去。

无论我们如何看待慈善本身，毫无疑问的是随着全球金融危机恶化、政府努力融集资金维持其社会福利体系，市场力量变得更加意气风发。他们意图重新建构发展合作，并通过商业框架对社会工作进行革命性改造，而商业框架本身又会反过来深刻影响我们所生活的社会类型。新的“行善者们”会像审视其他任何形式的金融投资一样对待社会事业，这也是为什么他们需要通过度量来评估潜力、可拓展性以及短期收益可能性，他们没有耐心等待，也没有时间等候一步一步进行的发展，他们希望一蹴而就，“这符合他们成果导向的价值观以及他们自己的行为方式”。他们对风险慈善的理解与推动社会、政治变革不可避免的风险毫无关系：风险慈善基于商业成功的确定性。在我们看见根深蒂固的不公正的地方，他们看见的只是这里缺乏有效率的方法而已；在我们看见权利失衡使人们遭遇贫困的地方，他们看见的只是这个地方需要创业机会而已；我们强调社会、政治赋权的重要性，他们只相信科技治国论下基于“硬”数据与经济模型的解决办法：“他们想要ROI（投资回报），SROI（社会投资回报），FROI（经济投资回报），以及EROI（情感投资回报）。”

这些慈善家都是实干者，不是社会科学家，对于他们，凡事都易如反掌。他

们已经建立了成功的企业（某些企业甚至比它的目标援助国家还要富有），比尔·盖茨会说：为什么不大胆解决一下世界上最严重的问题呢？“放手去做（Just do it）”，这是耐克公司人尽皆知的标语，无独有偶，耐克基金会与耐克创始人菲尔·奈特（Phil Knight）恰好都是首屈一指的慈善机构与慈善家。他们酷爱技术问题，因为这些问题清晰、易于辨识并且可测量。拿盖茨基金会支持的项目之一疫苗举例，他们的思路是：这里有某个问题（某种疾病），能够找到某种确切的诱因（某种病毒），那么就可以通过某种具体的干预措施（某种疫苗）来解决。但这就是全球低收入国家健康状况低下的根本原因吗？2005年刊登于医学评论刊物《柳叶刀》的一篇文章，针对盖茨基金会对医疗部门的干预提出了批评，该文章提出，“盖茨基金会狭隘地将健康理解为技术干预的成果，这脱离了经济、社会以及政治环境”。《柳叶刀》着眼于医疗技术与卫生健康改革的历史进程，认为人类寿命的延长及其健康始终取决于良好运转的、惠及广泛人群的公共卫生体系而非医疗技术（当然，并不质疑后者补充性的重要地位）。但是，公共卫生基础设施的建立需要耐心、长期的决心以及国家政策支持。而且，大量的冒险也是必要的，因为这可以杜绝那些影响社会公共卫生体系的各种利益（尤其是私人的卫生服务供应者的利益）。奥巴马政府在美国建立公共卫生体系的艰难生动地说明了其间涉及的巨大风险与既得利益。疫苗的推广相比之下更简单、更快速以及不需要从根本上重新思考我们的政治、经济体系，但同时，疫苗（尤其在未获得接种对象同意的情况下）仍是医学专家争论的对象。比如说，尽管单剂疫苗与无针疫苗的成本效益非常高，但它们的发展也许会减少对健康婴儿的探访次数，“而这对于检测婴儿的健康发育成长来说至关重要”。在有些情况下，预防例如痢疾等疾病的有效疫苗则会导致面临预算限制的政府认为普及干净自来水与卫生设施等问题“远没有那么重要”。

盖茨基金会坚持，为儿童接种疫苗是减少全球人口增长与改善气候变化的最佳方式（“好人俱乐部”认为这两大问题是时代面临的巨大挑战）。疫苗能够放缓全球人口增长与气候变化？盖茨基金会怎么证实这一违背直觉的观点呢？答案是硬数据，盖茨基金会已经建立了模型来说明儿童预期寿命的增加与更低的出生率正相关，这很有道理。常识与历史证据告诉我们良好的生活水平很有可能带来较低的出生率（尽管法国与德国这些国家已经能够在保持出生率回涨的情况下实现高生活水平）。但是，无论我们是否同意这一研究结果，我们能够赞成这是应对气候变化的措施吗？答案是绝对的否定。上一次我的调查显示，绝大多数的非洲家庭都不是排放吞噬我们大气层的温室气体的罪魁祸首：人口正在缩减的古

老西方国家才是。大多数低出生率国家的消费模式才与地球难以相容的，而非（绝大多数）非洲家庭的生活方式。当然，我并不是在质疑疫苗的重要性，但是简单地将这一微妙的议题与人口控制、抗击气候变化联系起来是不对的。恰恰相反，将我们不公平、不可持续的发展模式输出到这些国家只会使气候变化恶化，因为我们最后很有可能只是培养了数十亿的过度消费者。

当然，矛盾出现在各个层面上。盖茨基金会已经被以各种各样的理由批评过了，其中包括其母公司微软一直在倡导知识产权领域的严格管制，这导致非专利药品一直难以进入世界上最贫穷的国家。而且，微软咄咄逼人的政策及其软件领域的实际垄断地位也限制了新的社会企业发明出微软所欣赏的创新成果的能力。正如我们所知道的，软件领域的大多数创新成果（从新浏览器视窗到小应用程序）都是由程序设计人员之间大规模的开源协作实现的，这样的做法也是微软及其创始人极力反对的。有趣的是，盖茨的好友之一、首屈一指的慈善家以及亿万富翁沃伦·巴菲特，却反对在非营利性行业引入他所偏好的对企业行为进行评估的评级系统。当被消费者谴责放任他的公司伯克希尔·哈撒韦（Berkshire Hathaway）对违反环保、人权规定的商业进行投资时，巴菲特认为在社会、人权或是环保测量层面上对公司表现进行评级的做法也没什么好处，他不会在投资时考虑这些因素。慈善资本家们在对待成果与测量上的信条是如此一致。

爱德华认为，盖茨和他的同事在社会领域支持的硬数据技术评估（通常通过图像与数据计算来强化）“在质、量与可持续性上都显示出了很大程度上的失败”。而且，通过基于“客观数据”上的评级系统与投资模型，这些新的社会活动促成了一种“垃圾食品式”的参与文化，某个人可以简单地根据电脑屏幕来追踪非营利性股票市场的走向、调整社会投资。他将慈善资本家的方法与民间组织、社会运动的方法进行比较，后者的特点在于宽松的目标与长远的视野，他还坚持认为“商业指标与对成功的衡量使得人们更重视规模、增长与市场份额，而不是社会中人们的合作的质量以及培养起来的能力与组织”。“非营利性部门之所以能够存在就是因为它可以对社会问题进行资助与投资，而营利性部门则不能，因为这些问题无法被测量”，西雅图社会风险伙伴国际公司的执行董事保罗·休梅克（Paul Shoemaker）评论道，“我们的意图不是通过那样的方式让非营利性‘市场’变得高效，但是我们对非营利性部门与营利性部门进行了同样的效率测量。”当我们的重心从社会变革转移到了基于成果的核算，短期效率则替代长期有效性。这样一来，获得外部资助的非营利和非政府组织便会渐渐不再关注政治主张、不再为社

会变革而努力，这些都需要“深厚的资源与根据情况立刻改变策略的能力”。

有些人强调，将社会工作转变为一系列数据的、抽象的模型的过程是非常阴险的做法，因为这一过程会造成“市场化的慈善业带来的是系统的变化而不是稳定的错误印象”。其实恰恰相反，新援助事业通过基于随机实验构造净收益、成本收益比率与因果关系，建立了“一套错误地将世界封存在本体论永恒中的、关于稳定的含蓄的话语体系”。新援助事业由于强调数据与成果，显得像是在强化“那个导致贫穷、疾病与环境破坏的系统”。有趣的是，大多数新的大型慈善机构也在通过资助的方式来模糊市场与公民社会的边界，因为它们大多数的捐款都是以证券而非现金形式发放的。尽管这能够使受益者从金融高利润企业时代的“馈赠”中获得更多收益，然而在经济滑坡时，它们也承担着极高的风险。但是，更严重的是，这类捐款将接受者变为了金融市场的非自愿利益相关方。根据2013年《慈善纪事报》杂志的报道，在过去的几年中，由于金融界的混乱，美国最富有的个人的捐款数量在持续下降，并且进一步造成了社会领域的动荡与周期性趋势。

全球的新施予者们没兴趣问更深层次的问题。他们技术治国论的视角不允许他们探索系统性的问题，包括显而易见的问题：我们是否应该接受这一看似会导致地方性不平等与财富（资源）集中到少数人手中的全球经济范式？而且，拉姆达斯认为，随着北半球日渐显著的所谓“发展滑坡”（包括恶化的不平等、非可持续消费模式、青睐精英的金融系统、与生活方式有关的广泛健康问题），慈善资本主义者们“试图对能够将这一发展模式的奇思妙想带给全球人类、社区的努力、倡议进行投资”。所谓的西方国家对商品农业造成的食品不安全，对化石燃料过度依赖造成的环境、社会后果，基于系统性掠夺自然资源的经济模型造成的不稳定，以及企业利益铁腕下的周期性民主政治崩溃了解得越多，“他们就越坚定地想要和所谓的发展中国家分享自己成功的发展策略”。

由于数据倾向于将复杂现象分解为可测量的单位，所以它们往往藏身于系统性贫困、经济失衡与获得资源的不平等渠道的相互联系之间。又由于社会测量往往会将测量对象简单化，所以它最后会产生扭曲现实、误导政策的风险。因为这些度量方法本身的性质，它们会奖励那些容易、直截了当的方法，这些方法更易被数据捕捉、更易被标准化实验测试。尽管历史往往显示，可持续的、长期的社会变革只能通过参与、对抗与政治行动实现，但这些方法却将我们的注意力转移至技术性的解决方法。尽管数据倾向于强调商业的重要性及其寻找“解决方法”的

能力，现实却复杂得多，我们今天面临的绝大多数不公正都是根深蒂固的权力结构的产物。但是数据并不会揭露这一事实，它们只暴露出冰山之一角，因此巩固不平等现象的腐败的政治、经济结构则被深藏起来。



---

## 结语

# 重新思考数据，重新思考治理

---

“如果数据是体现宇宙神秘性的一个方面，我们则容易被使用数据的人控制。如果数据是人类掌控混乱的方法，它就会成为科技精英手段的一部分。”

——D. 鲍伊尔，《数据的暴政》，2001

我们要如何与数据争论？正如这本书所说的，数据的天性即揭露事实的本质，不必介意它是否真的这样做。我们都知道数据只是社会现象的部分表现，并且它通常被科学地或是政治地操纵，但是，它们的表现形式与设计都是围绕实证概念建立的——客观实证。当我们看见一个数据，我们便产生了确定性——事实信息。数据并不像文字，不需要解读，数据至今都是权威的来源，因为它揭露事实，而事实无可争议。（本书分享更多索搜@雅书）

当然，这并不是说，数据研究永远无法被质疑。在学术界，我们时常对数据进行争论，也讨论方法论并努力保证全面公开。我们的学生非常了解这一点，当他们展示自己的调查时，他们总是害怕这个不可避免的问题：“你从哪里得到的数据？”在我们发表文章时，我们通常会将免责声明包含在内，尤其当数据模型构成了我们分析的主干时。我们，作为学者，乐于警告读者我们的数据也许不够完整，以及我们的模型所基于的关键假设也许将真实情况过分简化了。通常我们会通过脚注、尾注、星号以及附录来处理这个问题，或许是因为我们知道那些内容只有极少数人才会读。这有点像广告宣传，在某个公司将新产品所有令人赞叹的特征大肆渲染后，他们会在难以看到的广告底端用小写字体补充上“适用条款及情况”。

总体上，目前用一些数据来证实学术分析是发表文章的前提。在社会科学领域最著名的学术期刊如今也不愿刊登任何连一个方程式、几个回归分析或一个因素分析都没有的文章，这是“定性”调查者的寒冬。不幸的是，对数据的追求不可避免地削弱了我们调查的分析深度，因为我们趋之若鹜地在数据集之间寻找关联与数据的意义，极少思考我们的数据的真实质量，并且检查数据质量并不代表我

们能够保证平均值是正确的或缺漏的数据不会影响最终结果，它意味着质疑数据作为所研究社会现象的良好描述符号的终极可信度。这些数据真的能够告诉我们关于社会的有价值的事吗？数据只是简单地描述社会现实还是已经变成了我们改造社会的规范工具？以及重要的是，我们有教我们的学生用批判眼光看待数据吗？不幸的是，人们对于自己知识的自满已经蔓延到了整个学术界，损害着我们的创新思维以及我们挣脱既定概念的能力。而且，免责声明不管听起来多么真诚，都不足以“将我们从邪恶中拯救出来”。

毫无疑问，数据已经不仅主导了学术思考，也主导了我们自己对于学术界作用的理解。首先，我们不断受到标准化的绩效评估。这一趋势最早起源于20世纪70年代的北美与英国，随后蔓延至欧洲和亚洲、南美洲、非洲的大多数新兴国家。学者们围绕着一个关键数据——H指数而生存，该数据是一种复杂的、用于计算我们的文章在其他学术文献中出现频率的方法。谷歌学术——数据革命时代的强大学术资源库——也已经成为施加在我们许多人身上的诅咒，因为大多数指数都是由谷歌自动计算得出的，而当谷歌没有注意到我们当中的某一篇文章时，我们就有麻烦了，“发表或湮灭”已经成为对我们的诅咒，“引用或被引用”则是我们新的铁律。我们学术生涯中另一项基本的指数是所谓的“影响因子”，它是由信息机构路透社计算得出的，显示了一本指定期刊中文献引用的平均数。如果你想在数据的学术丛林中生存下来，你必须努力在高影响力期刊上发表文章，剩下的期刊则无关紧要了。但是这些数据真的能够代表一个人研究成果的质量吗？或许能。但同时，它也会产生反常动力，因为数量变得至关重要，学者们都感到了达到标准要求的压力。数量下降时，全体研究人员都变得焦虑起来，他们开始召集特殊会议，职业生涯岌岌可危，名誉与耻辱已经成为司空见惯的事。学术人员之间互相引用已经司空见惯，自引比比皆是。审查者们通常只批准那些引用了一篇或多篇他们自己文献的研究发表，引用量已经成为晋升、涨薪或在常春藤大学谋职的关键，这些数据不仅对我们的名誉，也对我们的银行账户有着根本的“影响”。并且，在这场数据引发的狂热中，许多学者被曝光抄袭、重复发表以及伪造数据也是意料之中的事了，他们都极力试图跟上这种大流。

## 为数据辩护

本书无意质疑数据在知识发展与社会进步方面的重要性，也不否定数据在支持决策与政策上发挥的作用。我深知，如果没有数据，公共决策就会被直觉与肤浅的花言巧语所主宰——这也是为什么集权政府总是不发布常规数据或隐藏数据。科学革命正是由于对宇宙的宗教解释提出了疑问，才削弱了教堂在世俗事务上的力量并在现代一往无前。数据有过（并且能够）解放性的潜力：它们能够赋予人类力量，打击独裁者们；它们能够赋予年轻学生力量，让他们挑战教授们以及教授们“伪造的”数据；它们能够帮助环保组织将造成生态破坏的污染性企业送上法庭；它们可以帮助公民了解为谁投票以及支持什么政策。在学术界，数据也很重要，不管人们怎样批评H指数与影响因子这类发明，对卓越的追求总是要求某种形式的学术质量评估存在。没人想看到一个死气沉沉的学术界：教授们得不到任何声誉或经济利益上的激励，没有动力完成优质的调查与出色的教学。

公共政策也需要数据才能运作。正如物理学家开尔文勋爵（Lord Kelvin）所说：“如果你无法测量一个东西，你就无法改善它。”并且，如同人们一再引用的商业大师戴明（W.E. Deming）的警句：“你无法管理你不能测量的东西。”测量是人类生活基本的组成部分之一，我们每天都在测量，并且基于此做出决定。拒绝测量本身不仅仅是天真的，对于社会而言也是不可能的，这会导致人类天性重要部分的缺失。测量对于沟通也很重要：如果没有对规模的界定，便无法描述当地市场的情况；如果没有温度计，便不会有医生的准确诊断；如果没有长度单位，便不会有木匠的精确工艺。我们的教育、卫生与住房都依赖于测量，并且我们认为那是进步的标志。

但同时，我们也拒绝测量许多东西，并且我们有充分的理由。比如，我们认为测量艺术非常滑稽，因为艺术的价值无法被翻译成粗糙的数据。当然，我们经常为艺术品定价，但我们很难测量自由女神像或金字塔的美并评估它们哪一个更坚固、优美或珍贵。我们也拒绝对友谊的量化测量，有些人可能认为自己非常幸运，比别人有更多的脸书（Facebook）好友，但没人会认为好友的数量（无论是现实生活当中还是在网络上）代表着友谊的质量。恰恰相反，我们感到，如果我们想要根据某种数据参数来测量友谊，我们便会冒犯它，我们会通过测量友谊将它变成其他东西：一个被剥夺了首要的使它如此重要的人类情感的数据单位。相

似地，我们也拒绝测量爱，我们不会用数值范围来量化父母的慈爱或夫妻间的和谐。尽管一些经济学家与心理学家或许会采用度量来测量社会生活中这些“软性”因素，但是我们——作为整个社会——畏惧引进标准化的抚养或婚姻评估的想法。

在自然科学中，人们用数据来描述物理现象，并且数据也有好坏之分。历史学家表示，自然科学也难以免于被过度简化与范式转换。正相反，科学的历史进程是一个不断质疑主导模型（及其数据）的过程，人们意图用更好的理论替代它，而反过来，这些理论在科学革命的持续进程中也会被挑战。不论这些物理数据是好是坏，它们都难以影响原子的运动，但是当数据被应用于人类现象时就不一样了。数据在进入社会生活的时候，也会改造现实。比如说标准化评估，它不仅是分析表现绩效的工具，也是指导工具。如果重要的是数量，那么学者们就会努力发表每一项研究的成果，无论这些研究的质量有多值得怀疑；如果重要的是在脸书上拥有更多好友，那么用户就会尽最大努力积累人脉数量，而不是巩固和老朋友之间的亲密度；如果爱的量成为测量家庭美满程度的解释性标准，那么父母就会在可测量的方面，比如短途旅行的次数、玩具的数量以及陪伴孩子的时长，而不是在彼此相处的质量上花更多心思。我们不仅面临着被误导如何定义优先事务的风险，还面临着失去欣赏无形事物价值的能力。如果对德国政治理论家汉娜·阿伦特（Hannah Arendt）的话加以解释，我们会说将数据系统应用于社会生活的问题在于，“并不是说数据是错误的，而是它们本来可以反映事实”。

## 在优劣数据之外

从治理的角度出发，数据在影响政策制定上存在双重问题。在看完整本书后，我们会发现有些数据纯粹是“恶劣的”，全球数据调查的质量相差甚远，比如构成基于国民生产总值计算的国民收入账户的数据调查。在许多非洲国家，收入数据往往不完整，并且很大程度上依赖于当地统计员的估算，这导致了数据上各种各样的不一致。然而，政策制定者与国际捐赠者却每天都在使用这些“劣质数据”来测量结构性改革、发展援助与宏观经济政策的效力。美国经济分析局也承认美国的国民生产总值数据质量并不总是很高，尤其是在经济危机期间。而且，数据在被每个数据机构以不同的方式不断修改、调整以及再计算，这也使得国际数据对比比我们想象的要难得多。比如，2013年年中，经济分析局引进了一项国民生产总值账户的改革措施来计算在研发、娱乐、文化及其他艺术创意上的支出。在传统的角度，这些费用都被视为商业成本，如今则被当作固定投资，因此充分地被加入到国民收入当中。这也引发了批评的声音：美国政府极力抬高预计经济增长的百分比，意图压低国债占国民生产总值比例的增长。

我们也知道有许多数据集都被操纵或是完全编造的。比如，评级机构已经承认使用信用评级来迎合它们的客户，并且投资银行也对外债数据进行了操纵。学者们被曝光使用虚假数据，某些情况下他们的畅销书与整个职业生涯都是基于伪造证据之上的。我们也知道同行评审制度还存在非常多问题，许多期刊都落入了数据陷阱，包括具有最高影响因子的一些期刊，它们把事实上人造的数据当作无可争议的证据。如今计算机科学领域常见的一句话“无用数据输入，无用数据输出”，毫无疑问地，暗示着任何被输入数据模型的“垃圾数据”，都会被再次输出，如果无意义的信息被输入模型中，那么模型则会生产同样无意义的结果。但对于我们每一个人来说，这一过程看上去都是基于证据进行的。

数据也可以被扭曲服务于特定政治议程。环保怀疑主义者已经使用数据证明悍马比普锐斯更节能，而持相反立场的科学家也使用随机实验证明吸烟不会致癌。1954年，美国作家达雷尔·哈夫（Darrell Huff）出版了一本小书《统计陷阱》，该书后来成为所有怀疑数据的学生的经典读物。哈夫通过生动迷人的描述，指出了许许多多不实数据（比如通过删截图表来过度强调细微差异）或对数据的粗劣解读可能导致的操纵现象。正如我们所知道的，“党派利用实证武装自

己来抵抗知识的冲击，也用它为非理性披上理性与科学的外衣”。

然而，如果质量是数据唯一的问题，那么更严格的观察、公开程序与监管就可以轻易解决它，但不幸的是，即使数据没有被明显地操纵，它们也会误导决策，因为数据只关注可测量的事物，不可测量的事物都被忽视了，这不可避免地影响着我们工作的优先次序。根据笛卡尔对外在物（*res extensa*，测量事实）与思想物（*res cogitans*，认知事实）的区分，可测量的变成了唯一关键的事实，其余的则毫无用处与价值。再拿GDP举个例子，即使我们有能力改善其数据的质量（而且我们正在为此尝试），国民生产总值仍是市场产出的测量工具，而它绝对无法完整地描述整个经济体的全貌。正如我们所知道的，国民收入账户并不包括未经市场调节的交易，因此我们的政府就会使用国民生产总值作为经济表现的测量工具，以追求在牺牲非正式经济领域的代价上强化市场的政策，这些非正式经济领域包括家庭服务、礼物经济。并且，由于国民生产总值是基于市场价格计算的，因此未标价的东西便失去了价值。根据本书内容，这种对于价格的强调会产生畸形的刺激因素，一方面，我们的治理体系会倾向于那些忽视免费事物如自然资源的发展模式，另一方面，这种趋势会鼓励政策制定者将公共资源私人化、市场化，以使它们在国民生产中具备可生产性。在国民生产总值的治理框架中，只有被货币化的事物才具有价值，因此，为了严肃地对待自然环境，它必须被测量、被货币化。许多经济学家会说所有事物都可以被标价，即使市场不存在，价格也可以在简单的刺激下出现。对于他们当中的许多人，重新改善国民生产总值只需要调整技术以及更准确的模型而已。但是正如我们在整本书里所看到的，这个货币化“陷阱”会暴露出许许多多的不一致，实际上，“准确”已经成为一个形而上学的概念。我们要如何准确地测量大自然的价值呢？谁来测量大自然呢？依据的标准是什么呢？我们要选择人类中心主义的方法吗，这样大自然的价值量就由人类来决定了？还是说我们要采用整体的生态系统观，将自然世界的彼此联系纳入考虑范围？我们分析过了全部影响现有方法论的概念缺陷，从支付意愿调查到重置费用计算，全部的这些估值模型都建立在无力的假设之上，并且都暴露出基本的概念缺陷，当它们被应用于治理模型时，即便没有蓄意操纵与（系统出现的）利益冲突存在，它们也会变成危险的政策工具。

在商界也出现了“数据到底会造就金融成功还是抹杀创新”的激烈辩论。两位著名的学者，哈佛大学商学院的罗伯特·卡普兰（Robert Kaplan）与波特兰州立大学的托马斯·约翰逊（Thomas Johnson），就上演了这样的一场辩论。在20世纪80

年代，卡普兰与约翰逊共同撰写了畅销书《相关损失：管理会计的崛起与衰败》，该书认为尽管成本核算是19世纪驱动新企业扩张的关键动力，但是在全球化时代，这些测量工具却对商业造成了根本性的荼毒，因为它们将经理从公司的生产性部分中剥离出来，使他们更加依赖于抽象计算来预估成本、价格与投资收益。但在该书发行之后，这两位学者却在学术道路上渐行渐远，卡普兰支持使用各种经济测量模型来运营成功企业，而约翰逊却坚持亲身实践管理与人为判断。卡普兰在一系列成功管理书籍中推广使用作业成本法与平衡记分卡，它们可以帮助经理“从大量数据中提取真正重要的少数数据与结果”。相比之下，对于约翰逊而言，这些经济学主导的测量工具迫使经理通过定量数据而不是具体的商业知识领导公司。反过来，这一趋势促使“现代商业迷恋通过数据‘寻求利益’……不论这样的做法会对维系人类组织的基本关系造成什么样的损害”。

因此，数据的问题并不是简单的优劣数据之分，而更像是一个哲学问题。在某种意义上，这类似柏拉图与亚里士多德之间的争论，对于前者，数据是事实的象征，对于后者，数据是发展知识的工具。柏拉图遵循了毕达哥拉斯的观点，认为数据能够揭露世界的本质。相反，亚里士多德将世界看作一个混乱的研究对象，在这个世界上，数学推理只能够引导知识，但绝对无法代表更高层次的事实，对于他来说，数据无法揭露“形式”——与世俗意义上的“实质”相比，它们只是分析错综复杂的大自然的启发式工具，并不比它们企图描述的实质更完美。

此类将数据看作本质与将其看作工具的经典区分至今仍影响着我们的当代社会，并且总体上来说，柏拉图赢得了这场争论。确实如此，尽管数据不完美且存在片面性，但它们倾向于获得自己的生命，它们将自己从现实世界中提取出来，又生产出一种虚构的“超铀”，它们体现出实证，却不代表某种现象：因为它们变成了现象本身。既然“代表”从概念上说就是不完美的，数据便无法代表现实，而是必须预示现实。

数据这种内在的力量就扎根在我们追求事实的天性当中，这也就解释了为什么许许多多数据——不论优劣——成为塑造社会中自满与屈从的有利武器。尽管数据看起来像发展知识的工具，但至今它们已经抹杀了我们锻炼批判性思维的集体能力，它们带来了助长愚蠢的风险。一个建立在数据上的社会会“威胁它自身，因为它在肤浅的证实过程上投资了太多，而牺牲了其他形式的组织智慧”，伦敦政经学院的会计学教授迈克尔·鲍尔说道。这种现象通过数据基础上的审计系

统，变成了一种“博学的无知”。这些标准化的方法“支持抽象管理理念，却牺牲了其他的绩效评估理念”，并且“倾向于照顾那些从经济角度可以被测量、审计的事物”。



## 治理、数据与公共领域

在政治研究中，从治理的角度讨论决策过程已经是稀松平常的事。治理这两个字暗示着权力的分割，在当代社会中，国家政府已经成为协调者而非决策权力的垄断者，从企业到非政府组织，他们更进一步与全球机构与私人实体合作。通过研究治理，我们可以发现三个截然不同却又彼此联系的部门：国家、市场与公民社会。这一描述绝对不准确，这三个领域之间的界限或许没有我们认为的那样清晰，比如，公有（国有）领域与私有（市场）领域之间存在明显的重合，尤其在许多政策的设计、执行与管理过程中公私合作激增的情况下。相似地，营利性（市场）与非营利性（国家与公民社会）功能间的区别也持续遭遇“混合”组织和公共事业企业的挑战，在“混合”组织（比如社会企业）中以团结为基础的功能与营利性功能融合在一起，公共事业企业由政府控制但遵循市场原则。

随着这三个领域间的界限变得越来越模糊，它们不仅相互合作，也为了权力而竞争。由于治理机制为影响决策提供了新的机会，国家、市场与公民社会之间也形成了相互平衡的动态关系，但是冲突的风险始终存在，因此，治理可以被看作政治对抗的新地带，在这里不同的（尽管通常是相互补充的）原则与价值观相互摩擦或合作，以形成集体的决定。总的来说，治理就是文化、社会、经济与政治霸权构建的过程。

也许，这三个领域中最为复杂的是我们一般所说的公民社会。什么是公民社会呢？这个问题自古以来就是哲学争论的重要组成部分，在政治思想史上，我们会发现许多种相互补充的（有时是相互对抗的）公民社会概念。在古希腊经典政治思想中，“公民社会”描绘的是一个“良好社会”，即相对于其他形式的治理，塑造了城邦的礼仪、法规与公民参与国家治理的形式。对于亚里士多德，公民社会是相对于“野蛮人”的未开化世界的、通过自治组织起来的社会，之所以称它为公民社会，是因为礼仪。在罗马的共和传统中，公民社会是活跃公民构成的整体，这些活跃公民有规律地促进共和国的社会、文化、经济与政治繁荣。在现代政治哲学里，公民社会的概念随着个人自由与权利的发展焕然一新。对于约翰·洛克（John Locke）与亚当·弗格森（Adam Ferguson）这两位现代自由主义的奠基人，公民社会代表着现代有产阶级，该阶级在具有内在压迫倾向的国家内创造了自主自决的空间。对于这些思想家，公民社会在根本上是一个政治概念，是个体权利

对抗国家压迫的堡垒。

相似地，亚历西斯·德·托克维尔（Alexis de Tocqueville）将公民社会看作相对于政府的自我组织的场所，政府则被视为强迫的源头。这个自我组织的世界的特点在于其中的组织与网络，它们超越了庇护—被庇护基础上的传统社会关系。在这个概念里，公民社会为民主价值与抵抗压迫提供了土壤。而且，根据哈佛大学政治科学家罗伯特·帕特南（Robert Putnam）所阐述的，这种水平交互深入地促进了社会资本的扩散与生产。另一种由黑格尔思想发展而来的传统观念认为，公民社会是存在于国家与家庭之间的群体与实体构成的整体。这一观念将组织与群体视作根据国家命令在整个社会进行文化渗透的工具，并且深刻影响了马克思将公民社会视为资本系统（资产阶级）的核心。意大利学者安东尼奥·葛兰西

（Antonio Gramsci）再次阐述了马克思的观念，他认为公民社会是根据一致赞同的理念建立起来的霸权的领域，这一领域与国家统治权力的领域相对。他比马克思看到了公民社会更大的潜力，他提醒道，公民社会还能够为人民提供反抗资本主义统治结构所需的空間。对于社会人类学家欧内斯特·盖尔纳（Ernest Gellner）来说，公民社会是西方自由主义相对于其他政治意识形态的决定性特征。

在德国哲学家尤尔根·哈贝马斯（Jürgen Habermas）看来，公民社会应该被视为沟通行为的场域，也就是所谓的“公共空间”，在这里人们讨论并处理想法与价值观。在这个观念中，公民社会为对话社会奠定了基础。这个空间处于国家与私人领域之间：人们在这里进行公共辩论、交换信息，群体与个体得以表达他们的看法与兴趣、讨论共同目标以及正视他们的偏见。这样看来，在公民社会中，话语互动——即对抗、争论与思考的持续过程——维系着民主的本质。公共空间不是静止的，它是一个不断发展的动态，在其中，对抗带来变革，反过来，这也影响着参与者的认同与价值观，因此重新刻画着公民社会本身的轮廓。公民社会作为公共空间的概念固有地假设了想法、目标冲突的存在。在高度统一的社会里，沟通、对抗与辩论无法出现，正如葛兰西在他对霸权的分析中所暗示的，公民社会作为一个社会空间，其特征在于固有的改造潜能。

无论人们赞同强调公民社会与国家、市场之间对抗的经典理论家们，还是赞同那些将公民社会视为沟通与思考场域的人，显而易见的是，在如今数据风靡的世界里，公民社会作为变革发生的场域这一政治本质正逐渐遭遇挑战。数据凭借其本质导致了辩论的减少，它们并非动态的实体，相反，它们的本质是静止的。

根据第一章，数据是被设计用于抹杀现代资本主义社会中的自由裁量权与避免政治争论的，它们是控制的工具，尽管并不给人以这样的印象；它们是统治的工具，但却不施加压迫。在福柯的话里，它们是政府统治的工具。数据的微妙之处在于它们并不消灭权力，而是将它隐藏起来，而正是这样的隐形能力，让它们在政治中如此有影响力、对公共辩论如此危险。

数据在公共政策中所表现出的形式：统计信息、评级、成果、测量与所有其他形式，都产生了“治理去政治化”的效果，因此使公民社会丧失了真正的改造潜能。数据通过将不同价值观、原则与想法扁平化为测量方法，从根本上改变了国家、市场与公民社会之间的政治交互。持续的想法的对抗本应该成为公共空间的特点，如今它们已经变成了决策中对效率的单维需求。数据的效率已经取代公共辩论的政治本质，成为寻找世界问题最佳解决方法的认知工具。一方面，这样的趋势使政策制定高度集中化，越来越多的权力流入数据的“主人”——即所谓的专家与技术官僚手中。另一方面，由于市场被视为数据推理的最理想场所，因此这一趋势已经刺激了新形式市场霸权的出现，这种霸权的特点是主导当代社会的狭隘的经济思想。正如我最后在《GDP究竟是个什么玩意儿》中讨论过的，国民生产总值的发明帮助人们创造出了有史以来最有力的说辞，即市场是财富的唯一生产者，以及无尽的市场产出是政治的终极目标。国民生产总值还使所有的污染工业“焕然一新”，它们当中的“次品”都被淘汰掉了，同时，在屈从为上的国家里，它们还征服了工会，也削弱了环保运动，使其无力就经济增长的必要性与可行性推动有意义的辩论。无论是地方的还是国家的治理机构，都是围绕强大数据建立起来的，而且，国民生产总值为央行行长、经济顾问、发展顾问、国际货币基金组织专家、世界银行的人等提供了巨大的权力，因为这些人最了解如何促进经济增长、管理商业周期。科技治国主义的力量在信用评级领域变得越来越显著，该领域内少数公司已经统治了整个地球，它们的数据操控着全球的政策，迫使整个国家进入紧缩状态，对社会正义与集体利益都造成了损害。再一次地，数据被用来加强市场对社会生活的其他领域的控制，同时为数据的控制者提供了巨大的权力。

如今，几乎所有的社会、环境政策都要接受成本效益分析的评估。这些方法论一开始看起来非常理性，在采用一项政策前测量其成本与收益很有道理，但是再看仔细一些，我们会在其中发现不合理的假设与存在偏见的定义。贴现率系统地将更高的价值分配给了当下而牺牲了对未来的估值，因此产生了明确的采用短

期方法的偏见。而且，将成本与收益合并成为一项数据也隐藏了这样一个根本的问题：是谁的成本和谁的收益？在社会中，成本与收益并不是由全体成员平均分享的。在涉及监管时（尤其在环保领域），成本由企业承担，收益最后由全社会共享，但是成本效益分析过度地表述着“他们的”成本与“我们的”收益，好像这两者处于同一水平，导致所谓的高性价比解决方法比起基于社会公正与环境可持续性等原则的解决方法更受青睐。这样的情况在过去几十年的主要环保法规当中都一直存在，在抗击气候变化的治理中市场的胜出就是最好的例证，其中包含的思想是：气候变化可以通过正确的公式、寻找关键的均衡值与最佳定价机制来解决。但是根据许多分析人员的观察，“气候变化是无法用优雅的方法来解决的，因为它是一个分散的问题，更好的理解是，它是一种特定发展路径的症状”，这种症状构成了“一个复杂的关系网络，其中包含着相互强化、错综复杂的人类行为模式、物理材料与生产成果的科技，想要通过仅仅关注单一的方面把如此复杂的系统改造成理想的样子是不可能的”。

根据这本书的内容，不论市场失灵多少次（比如在信用评级与碳排放交易问题上），数据的力量都能够毫无疑问地让所有看似应该破产的想法与方法起死回生，“慈善资本家”逐渐强大的力量就是该领域的清晰例证。在慈善资本领域，测量驱动的企业经营思想无论在资源还是操作方式上都占有支配地位。爱德华认为，公民社会“能够在其生态系统保持健康多元的情况下呈现出最佳状态”，但是由于慈善资本家们所赞同的测量驱动的意识形态，“中介团体、非政府组织与他们应该代表的选民间的距离正在拉大”，并且“那些用于团结不同阶级、地域与种族的古老组织正在消失”。2003年由凯洛格基金会（Kellogg Foundation）发布的一项报告认为，“对效率与市场份额的强调会危机非营利部门的最基础价值——在社会内部创造‘自由空间’，使人们可以为社会问题与公共福利找到解决方法”。

我们在生态多样性保护领域也看到了市场影响的蔓延。由于测量自然资本已经成为一项有利可图的生意，金融公司持续发明出新的测量工具、审计与其他定量方法来为大自然标价，但这些数据背后隐藏的，是一个充满利益冲突、投机交易与其他有害倾向的世界，数据模型显而易见的客观性让这一切都隐形了。通过数据治理系统，我们仿佛将大自然当作了金融交易中的对手，我们允许所谓的专家来设立阈值、评估风险，然后，在这些数据的基础上，我们又与地球母亲玩着危险的游戏：我们根据生态破坏、环境风险与气候变化的评估情况来忖度自己的行为，而不是鼓励通过开放辩论来讨论人类走什么样的发展道路才能够带来社

会、经济与环境福利，我们用许许多多可疑的模型来与大自然“赌博”，企图花最少的力气实现最大的收益。

## 现在怎么办？

当下的治理模式正通过模糊数据政治来扼杀就损害我们社会的最深刻的不平等现象所进行的理性公共辩论，有些人认为，依据数据进行的治理会“剥夺我们为自己考虑的民主权利”。测量、评级、成本效益分析、标准化评估、投资收益与定价模型的系统应用导致了人们所说的“市场的必然性”。由于我们的概念工具受数据推理影响太大，因此我们的思考也被禁锢于市场的“匣子”内。如果我们对耶鲁大学经济学家查尔斯·林德布卢姆（Charles Lindblom）的话加以阐释，我们也许会说市场已经成为“概念监狱”。不论市场失灵多少次，不论我们多少次察觉到市场哲学固有的短期主义，我们似乎都注定要服从于市场对我们都得主宰，不仅对我们的治理系统，也对我们的智力。

我们的想法已经被数据禁锢了，这也是为什么我们甚至会对数据的失败投以更多的信任。在1987年发表的一篇文章当中，社会学家苏珊·夏皮罗（Susan Shapiro）对她所称的“社会对非个人信任的控制”进行了分析，她反思了“谁来保卫我们的守护者”的问题，并且提醒道，社会假定其守护者会“说出真相、良好地履行义务、遵循既定程序并且如同无私的受托人一样采取行动”。但是当他们辜负了我们，我们却不会重新思考政治、社会机制，而是“在糟糕的投资后又抛出‘好的’钱”：也就是说，我们“保护信任与应对市场失灵的方式是付出更多的信任”。戴维·博伊尔（David Boyle）说过，“现代社会的悖论之一就是，审计失败的解决方式是雇佣更多的审计员。”我们为什么要这么做？或许是因为，我们作为社会动物除了相信没有其他选择，社会最终是建立在人与人的相互信任上的，而我却对这个答案不满意。事实上，不信任在我们生活的世界大行其道：我们用防盗门来保护自己的家；我们很少与邻居往来；我们希望合同都包含责任条款；我们在街上走的时候会紧紧抓住钱包。但神奇的是，当我们讨论市场治理模型的时候，这些不信任统统蒸发了。在这种情况下，我们有意识或无意识地相信了评级、碳交易市场、补偿机制与生物多样性信用的奇思妙想。尽管我们会仔细检查街上的摊贩有没有给我们找对钱，但我们却深信不疑地假定，由经济学家团队操作的成本效益分析或由专家执行的投资社会收益是值得信赖的。

这本书就这一悖论给出的答案是，数据将治理领域变成了技术性过程，这个过程投射出一幅专门技能与专业主义的景象，而且，数据将治理过程变成了可以

依靠适当技术管理的任务。的确，这个想法让人感到慰藉，它将世界简单化了，暗示这个世界上对应分散的问题存在分散的解决办法，最重要的是，它显示只要依据正确的公式，我们就可以做出惠及每个人的决定：你和我、穷人和富人、污染者和地球母亲。在这个日渐恶化的地球上，我相信听到最严重的问题——从金融危机到气候变化再到全球贫困都有简单的解决办法，能够让每个人都感到安心。我们将挽救世界的任务授予那些数据的主人，这样一来，我们的终极目标似乎就变成了享受每一天被剥夺权力后的轻松。

尽管这种授权感非常令人慰藉，但现实是我们生活在一个复杂的世界里。虽然数据对于人类发展十分重要，但它却是一把双刃剑，它会不知不觉地降低社会现象的复杂程度，最终将我们带向错误的方向。就像一位尽职尽责的母亲绝不会将她的角色总结成一道公式，我们也不应该期望经济测量模型能够自动地操纵治理系统。治理是一项公共福利：最重要的公共福利。不仅人类的未来，而且整个地球的未来都取决于我们从全球到地方的各个方面的治理承诺。公共空间在市场理性主义逐渐增加的压力下收缩得越多，我们越难恢复控制民主制度的能力。更危险的是，由于市场机制排挤掉了其他形式的社会交互，我们也会彻底消灭其他替代形式的社会化，因为它们无法在传统角度下被测量。礼物经济、社区互惠方案与其他类型的非正式动态在正式市场结构的压力下逐渐消失，在这个过程中，我们不仅失去了整个社区与生态系统，也失去了数千年来的知识。

我们吸收着基于测量的“博学的无知”，却抛弃了其他的生活方式。我们变成了单维的人类：越来越符合应用于他们身上的数据模型的效用至上者。2009年，埃莉诺·奥斯特罗姆（Elinor Ostrom）凭借她关于公共资源管理的著作获得了诺贝尔经济学奖。她接受了作为政治科学家的训练，有勇气反抗数个世纪来的传统经济学知识，她认为治理的道路不止一种，私有化与商品化，或是自上而下的监管，都并非人类管理公共资源的唯一途径。她周游世界，从日本到瑞士、从加利福尼亚到菲律宾、从加拿大到土耳其，向人们阐述着自下而上的集体治理系统。在这个系统中，公民们建立起共享的制度与集体的合作性机制，不仅为了实现更好的，也是更坚韧、更平衡与更持久的治理成果，这就是为什么公共空间如此重要。全部的这些社会生活软因素，从互相尊重到团结一致，都无法被系统地测量，但最终都比那些被纳入推动当代治理的数据模型中的因素更加重要。诚然，参与也许是一项惨痛的经验，交互、辩论、妥协与思考也许令人疲惫沮丧，但除此之外别无他法。我们都是社会动物，生活在深刻交互的世界里，拉杰·帕特尔

（Raj Patel）评论道，我们无法从市场社会中找到解决方法，它只存在于“共同生存与参与使我们珍惜共同未来的民主政治的自由当中”。数据无法拯救我们，我们只能依靠自己。



---

## 致谢

---

在《GDP究竟是个什么玩意儿：GDP的历史及其背后的政治利益》一书出版后，我收到了来自世界各地的邀请，与大家一同探讨书中观点。由此，我开始了一段精彩的旅程，并在其中遇到了各行各业的人：学者、媒体人、广播主持和关心社会的公民。他们致力于深刻的、可持续的和激进的变革，他们提出问题，热衷于挑战传统智慧并寻找新想法来提升工作。所以，我要首先感谢他们，因为他们是重要的灵感来源。

在海德堡大学社会投资中心的格奥尔格·米尔登贝格尔（Georg Mildenberger）和埃克哈德·蒂穆勒（Ekkehard Thümler）的帮助下，我对全球治理中公民社会和市场的作用做了很多思考，而比勒陀利亚大学的马克西·斯库曼（Maxi Schoeman）、穆吉希·科布（Mzukisi Qobo）和卡米拉·阿德尔（Camilla Adelle）在这场智慧之旅中与我相伴，我们共同参与了治理创新研究中心的建立，本书中大多数的研究话题最终都在此获得了进展。我还要感谢我的各位同事，他们与我分享了对经济绩效、可持续性和市场支配力衡量问题的关注，特别是新经济基金会的萨玛·阿德拉（Saamah Abdallah）和托尼·格里纳姆（Tony Greenham）；负责Demos可持续发展倡议的卢·戴利（Lew Daly）；转型战略的克劳迪乌斯·范维克（Claudius Van Wyk）；斯坦陵布什大学可持续发展研究所所长马克·斯威林（Mark Swilling）；夸祖鲁·纳塔尔大学公民社会中心的帕特里克·邦德（Patrick Bond）；以及西德尼·勒基特（Sidney Luckett）。还有，非常感谢苏珊·乔治（Susan George）的努力贡献、深厚友情与学术指导。

我在Zed Books的编辑肯·巴洛（Ken Barlow）说服了我，让我以创纪录的时间完成了本书，他相信这将对全球数据应用问题的完美研讨。他耐心地审阅了我的手稿，纠正了一些错误，尽管书中涵盖的话题非常复杂，他仍然尽力帮助我提高了文本的可读性。我的妻子雅尼娜（Janine）一直是书稿的忠实伙伴和细心读者。无意冒犯我的同事们，但我必须说，妻子是我所知的最棒的智者。

最后，我要感谢我的学生，他们让我参与了许多富有成果且具有挑战性的对话，让我意识到在教学中有激情是多么重要的一件事。他们充满着想法，但传统

的教学方法总会压制他们的能量。我希望这本书和以前一样，能进一步激发他们的自由思考。

如果您在此书中发现了任何错误，我对此表达诚挚的歉意。