arXiv:2611.27923v1 [cond-mat.stat-mech] 20 Nov 2026

绿色强制措施的财政与经济影响：公共治理的玻尔兹曼方法

Katerina Savvaidou[1] 和 **Georgia Foudoukidou-Louka**[2]

欧盟财政政策与行政让·莫内讲席 法学院，塞萨洛尼基亚里士多德大学（*AUTH*）大学校园，希腊 *54124*

and

Ho-Ryun Chung[3] 与**Orestis Loukas**[4]

医学生物信息学与生物统计学研究所 马尔堡菲利普大学 汉斯·梅尔魏因街*6*号，德国 *35032*

摘要

公共当局日益需要将多方面的政策目标转化为具体的财政与投资选择，同时兼顾其社会经济后果。例如，在欧洲经济与财政治理中，绿色预算为将环境目标转化为预算政策的具体目标提供了一个框架。在实施具体政策干预之前，评估实现此类目标所产生的影响，需要进行可信的事前分析。

统计模型和数据处理流程的多样性引发了一个问题：仅凭历史基线（通过基础设施项目库体现）和政策目标，是否足以得出唯一且易于处理的评估结果？为了解决这一根本问题，在缺乏有关项目选择的更多具体知识时，我们将统计物理学的大规模视角与此前关于粗粒化抽样概率的研究成果结合起来，并应用于公共治理。

对于任何可通过少数几个关于众多不同项目的汇总条件表达的政策目标，我们借助条件抽样下的认知一致性，构建出一个理想化政策情景，使其与历史基线的偏离最小。在玻尔兹曼表述中，这一政策情景表现为一个反事实库，其项目构成精确满足指定的政策目标，同时与历史存储库的偏离最小。对于规模足够大的历史存储库，项目选择的细节实际上可忽略不计，因此这一反事实库中的任何维度都可轻易归因于历史基线和指定的政策目标。

[1] savvaidou@law.auth.gr [2] georgiafl@law.auth.gr
[3] ho.chung@uni-marburg.de [4] orestis.loukas@uni-marburg.de

在最小偏离政策情景下，可基于历史存储库中任何可计算的度量型，研究实现指定目标所带来的下游影响。在同一政策情景下，通过对代表替代历史版本的项目库抽样概率进行粗粒化而构造的内在p值，可对这些目标与历史基线的兼容性进行统计评估。就历史项目数量的主导阶而言，所有下游量都可通过闭式表达式高效估算。

作为原理验证应用，我们分析了世界银行数据库中的发电项目。我们的框架揭示了在可行的绿色电力盈余范围内，预期项目成本呈现非单调变化特征；通过资本效率系数识别出推动绿色转型的技术驱动因素，并揭示了资金分配和资本效率方面现有区域差距持续存在的风险。与此同时，历史绿色电力缺口在统计上仍与任何水平的可行绿色电力盈余不相容。

关键词： 统计物理学；条件抽样；普适性；公共治理；事前影响评估；下游分析；绿色预算；政策情景；玻尔兹曼组合学；粗粒度抽样概率；认识论依据

目录

1 引言 3

1.1 政策背景与方法论要求 .3 1.2 事前评估的传统方法 .4 1.3 统计物理学视角 .5

2 历史数据、项目概况与绿色分类 8

2.1 从原始数据到可观测量与项目概况 .9 2.2 绿色分类与政策可观测量 .10 2.3 次级可观测量与完整项目概况 .12 2.4 历史项目库 .13

3 玻尔兹曼抽样的反事实世界 14

3.1 绿色强制要求的假设.16 3.2 玻尔兹曼假设分布 .19 3.3 操作层面的影响 .23

4 结果 26

4.1 绿色转型的统计评估与财政影响 . . . . . 26 4.2 绿色转型的进一步影响 .31

5 结论与展望 **36**

5.1 结论 .365.2 展望.38

A 附录 **42**

A.1 玻尔兹曼统计框架概述 .42A.2 数学背景 .45A.3 进一步的下游分析 .51

1 引言

1.1 政策背景与方法论要求

气候授权通常要求将各种资源重新分配，以服务于特定的政策目标。希望加快向更绿色基础设施转型的政策制定者，不仅需要评估该目标能否实现，还需考量其实施对投资分配和项目成本、技术组合以及进一步的地理和经济方面会产生何种影响。在欧盟经济与财政治理中，绿色预算为将环境目标纳入预算决策并评估其更广泛的财政和社会经济影响提供了框架。尤其是，气候目标的财政影响可能因所采用的政策工具而有很大差异，具体体现在公共支出、私人投资、补贴和收入侧措施之间的平衡上 [1]。

事实上，大多数政策目标本质上都是多维的。在政策制定者能够获得现有项目库的历史基线，而该项目库包含公共计划及其支出，或包含项目及其投资时，一系列事前问题便会随之出现。在当前的技术能力和预算约束下，政策目标是否能够实现？实现该目标将对不同技术和区域之间的资本重新配置产生何种影响？私人投资以及其他并非由政策制定者直接控制的维度将会发生什么变化？此外，在进行任何重组之前，现有的制度和市场环境——正如历史基线所反映的那样——是否已经与政策目标合理兼容？需要注意的是，这些问题关注的是实现强制目标所产生的影响，而不是对某个假定的潜在总体模型进行估计。

为回答此类问题，政策制定者会在事前影响评估中考察强制目标所蕴含的下游后果。该评估先于实际干预的战略规划，并为其提供依据，而实际干预旨在实现这一目标。这些影响可能包括财政支出、经济指标、技术构成、区域分布或其他相关指标的变化。政策制定者可以将这些影响与适用的财政、经济和政策约束进行比较，以评估计划中的干预是否符合现行框架。为此，公共治理原则，包括比例原则和健全的财务管理，构成了

评估下游后果的规模和性质时所依据的规范基础。

为了使所报告的后果具有意义且便于解释，其来源必须能够从评估本身中清晰辨识。欧盟和经济合作与发展组织监管治理中成熟的、基于证据的事前影响评估框架强调，政策评估应以明确规定的政策目标、可获得的证据和适当的基准为基础，并应考察与这些目标相关的备择假设行动方案的后果 [2, 3]。然而，即使遵循这些指导原则，也无法确定唯一的定量评估，因为相同的历史证据和政策目标可能支持不同的分析，具体取决于设计时所作的选择。这种裁量空间使分析者、专家和其他利益相关者能够对评估结果施加重大影响，并进而影响由此得出的结论 [4]。更根本地说，对同一政策目标作出不同的评估，提出了一个实际问题：对于同一度量型指标，如何从多个估计值中选择一个；例如，参见 [5]中的相关讨论。

为了促进民主协商中的多元主义，我们需要针对明确规定的问题给出一个定量答案——不是为了终结争论，而是为了提供共同基础，使有意义的分歧和知情决策得以产生。在从可获得的基准中指定一个历史项目库后，任何政策目标的事前分析都应明确规定目标、相关结构条件，以及至多一组与公共治理相关、由独立主体采纳的原则。事前影响评估应能够清晰地分离这些核心输入对所报告后果的贡献，并将其与分析者、专家或其他第三方所作的选择区分开来。我们将这一方法论要求称为可归因性。

接下来，我们将探讨较为传统的统计方法是否以及如何处理可归因性，然后提出对政策问题的一种备择统计（重新）表述。尽管我们将所提出的框架应用于绿色转型的气候授权，其底层逻辑并非气候政策所独有。我们的框架很容易推广至公共治理其他领域中政策任务的事前统计评估及下游分析。

1.2 事前评估的传统方法

传统的定量方法借助统计推断、预测建模、工程模型，或这些方法的组合。尽管这些方法可服务于重要的科学和预测目的，但其设计未必与本文所探讨的事前政策问题相契合。后者关注的是，在基于双方认可的历史存储库的前提下，实现既定政策目标所带来的影响。因此，评估的核心是可归因性，而非推断底层的数据生成过程或预测未来结果。此外，对此类问题相关的诸多现实世界影响进行建模，不仅会导致计算上难以处理，还会使对问题作出全面描述在认识论上成为不可能。

然而，当以模型为中心的方法被用于这项任务时，必然会引入简化假设，例如结构性约束条件和依赖关系

而这些假设并非由政策问题本身直接决定，也没有公共治理原则作为依据。统计建模要求我们确定变量的边缘分布和联合分布，以及变量之间的关系，即使任务本身几乎没有促使我们作出这些决定的理由。关于模型不确定性的文献早已认识到，建模选择会对政策结论产生影响 [6, 7]。既有指南建议开展敏感性分析和不确定性分析，包括比较不同模型，以探究这些选择所带来的影响 [3, 8]。

似乎不存在一种系统化、以模型为中心的方法，能够从根本上消除政策问题范围之外的其他构念对所报告的下游后果及统计合理性度量（例如p值）的影响。人们或许希望这些构念对所报告数值的影响仍然很小，但这无法在事前分析本身中从根本上得到证实。在最坏的情况下，建模的影响可能会严重损害可归因性，使规定目标所报告的后果受到与政策制定者目标无关的构念的显著影响。

例如，逻辑回归是一种跨学科广泛使用的标准方法，常被用作构建更复杂模型的基础。从最大熵的视角来看，逻辑模型纳入历史项目库中响应变量与预测变量之间观察到的成对关联，并将高阶效应限制在这些成对结构所隐含的范围内。仅遵循历史基线中的成对统计数据，或许是出于建模稳定性或预测准确性的考虑。然而，从政策制定者的角度来看，究竟是哪项政策目标、结构条件或治理原则要求施加这种特定限制，仍不清楚。可归因性方面的疑虑恰恰在于，建模所得的三阶及更高阶效应通常会不同于历史项目库中的相应效应，即使规定目标在历史项目库中已得到满足。至关重要的是，对于涉及这些高阶效应的下游指标，该模型没有内在的方法来判定由此产生的差异究竟微不足道还是影响重大。因此，政策制定者无法确定所报告的后果是否确如预期那样由明确规定的政策目标所隐含，还是已受到模型结构的重大改变。因此，即使应用过程透明、数学上严谨，这些方法也可能严重不适用于事前评估。

顺便一提，人工智能可解释性问题也源于同样的顾虑。让评估的运作过程易于理解，并不能自动证明其结论所依赖的既定假设是合理的。将可归因性作为事前分析的内在要求加以保障，促使我们提出一种备择表述：从历史存储库出发，在明确可设定的约束条件下纳入政策目标，同时不对底层过程引入临时假设。

1.3 统计物理学视角

如上所述，大多数传统统计方法都面临一个基础性的建模难题，但这并不意味着事前政策问题必须无法进行定量分析。资料库形成过程复杂，受到社会、经济、技术、物理和制度因素相互作用的影响，而这些因素的综合影响

难以厘清，因此需要一种备择假设的数学描述。这种描述应允许在无需穷尽解释其底层机制的情况下进行定量分析，同时以政策问题直接指定的评估维度为依据。

一种可能的思路源于19世纪和20世纪的统计物理学。统计力学提供了一种系统的方法，用于描述更基本单元的集体行为及其大尺度组织，而无须详细描述每一种微观状态。尽管这种大尺度行为是由许多底层机制共同作用而涌现的，但其定量描述通常在数学上易于处理，而且无需逐一明确说明这些机制，也不必对它们之间的相互作用进行显式建模。将统计物理学思想应用于社会经济与政策问题时，通常以相互作用的主体为基本单元，利用统计力学结构描述个体决策及其相互作用如何产生集体现象 [9–11]。

统计物理学视角确实为描述集体现象中的涌现规律提供了一种有价值的方法 [12, 13]。与任何集体行为描述一样，其适用性取决于单元数量是否足够大。在我们的政策情境中，如果历史存储库中只有少数项目，就需要详细了解其形成过程，包括促成这些项目被选定和实施的决策与具体情况。如果无法获得这类信息，通用的统计描述便无法取而代之。引入更复杂的模型不仅无法自动解决缺失信息的问题，还会带来不可控失真的风险。

当历史存储库中包含的项目足够多，以至于可以将它们的选择视为近似独立的选择过程时，情况就会发生显著变化。在这种情况下，所有现实世界机制对资料库形成的贡献（例如，可以通过相互作用的主体图景或任何其他可行的、更基础过程的模型来理解）都可以汇总描述，而无需逐一重构或建模每一种现实世界过程。带着这一视角，为什么不把项目库中的项目用作形式描述的构成要素呢？毕竟，政策问题通常关注的正是这些项目。

当历史项目库包含大量项目N 时，我们将有效地把许多现实世界影响对项目选择造成的不确定性视为“运气”或“偶然性”。通过一个思想实验（思想实验）可以最容易地说明这一点：在该实验中，我们反复从一个瓮中抽取项目。这一思想实验描述了我们如今所称的条件多项抽样，其普遍相关性最早由 L. Boltzmann 于19世纪后期阐明 [14]。因此，从瓮中进行一般抽样，将集体行为在适当的大N 情形下所促成的项目选择有效统计描述概念化了。[1]

[1] 不确定性可以通过不同的抽样方案来表示，例如超几何分布或泊松混合；在适当的极限条件下，它们可以导出相同的有效统计描述。这些概念化方式都不是为了重现历史存储库形成过程中实际发生的事情。它们只是提供了不同的思考方式，帮助我们理解在有效描述中普遍呈现的同一统计结构。

值得注意的是，当上述大N 条件适用时，这种直接以不同项目概况为基础构建的简化描述，可以像明确考虑所有潜在机制的描述一样，准确捕捉聚合量正在呈现的行为，例如与刚才所述政策问题相关的平均项目成本。根据统计物理学的核心直觉，当许多影响因素共同作用于某个聚合量时，只要不存在阻止大N 简化的系统性机制，对其行为的描述便无需考虑这些因素的细节（这些细节或许能够解释单个项目是如何形成的）。如果这种系统性扭曲确实存在但不为人知，任何统计描述都会受到损害。

当许多单位共同贡献于我们关注的量时，概念上的简化并不自动保证这些量能够得到计算上可处理的评估，即使采用有效的多项抽样方案（从无限 urn 中独立抽样）也是如此。幸运的是，政策问题通常可以自然地表述为关于广义矩的聚合条件；广义矩定义为物理、财政或经济可观测量函数的期望值。对于这类将广义矩限制在指定值的问题，我们先前的工作 [15]已表明，粗粒化的抽样概率质量（包括内禀p值）在N的领先阶上具有半解析的闭式表达式。要使推导保持有效，聚合条件的数量必须远小于历史项目库中的项目数量，而在实践中，这几乎可以肯定成立。

基于以一般抽样方式构建的有效描述所具备的可处理性，我们利用内禀p值的认知一致性及其对经验结构的可归因性，推导出最小偏离反事实情景。由此产生的I投影支配着抽样概率的鞍点近似，它描述了一个理想化、无限大的项目库；该项目库相对于历史基线的偏离程度最小（以信息散度衡量），同时又足以满足政策目标。在有效描述的说明性思想实验中，这个反事实库代表了政策目标所隐含的理想化全集。

因此，利用一般抽样在大N 极限下呈现的内在结构，我们从形式上和实践中证明：给定历史存储库，政策问题本身会明确无误地构建事前评估的反事实情景。按照构造，这些最基本的要素会自动确保在适当的大N 状态下具备可归因性。这种简约性与我们在第1.2节所述、针对同一问题的以模型为中心的方法所具有的复杂性和间接性形成鲜明对比。

利用构建的反事实库，我们可以进行统计评估和下游事前评估。我们既可以评估历史基线与规定目标的统计兼容性，也可以通过历史项目库中任何可计算的度量型指标，评估该规定目标带来的后果。闭式表达式（由鞍点附近的高斯化得出）与条件I-散度最小化的高计算效率相结合，使这一框架适用于反复应用：随着历史项目库更新、政策目标演变以及新的事前评估开展，该框架均可持续使用。事实上，随着历史项目库不断扩大，有效的大N 描述

预计将变得越来越准确。相比之下，以模型为中心的方法可能需要随着项目库扩大而进一步对高阶效应进行建模，从而增加模型复杂度。

除计算效率之外，该框架以统计物理学而非公理化统计表述为基础，因而易于扩展到更一般的结构条件和政策条件。另一条自然的探索路径是广义抽样方案，这些方案纳入实际项目库形成过程中的程序细节，同时保留可归因性保证。

论文提纲

本文从基础设施项目历史库的界定出发，进而提出并评估反事实政策情景，并以绿色转型为背景加以例示。首先，我们在第2节描述数据准备以及合理可观测量的选取。这些可观测量界定了不同的项目概况，构成所提出的面向公共治理的玻尔兹曼框架的基本要素。

随后，我们阐述与绿色指令相对应的组合假设，并辅以直观的结构条件，见第3.1节。基于大N 结构下的粗粒化抽样概率，玻尔兹曼假设分布的选取在第3.2节中得到论证，并在附录A.2中作更详细说明。玻尔兹曼假设反事实对公共治理中统计评估的操作层面影响，于第3.3节加以探讨。

根据世界银行数据库，第4 节和附录A.3 通过图表和图示，展示了我们对所有可行绿色电力盈余的事前影响评估。我们考察了规定目标与历史基线的统计兼容性，以及在最小偏离情景下实现这些目标所产生的影响，涵盖经济、技术、地理和投资相关维度。

在第5.1节中，我们总结了全面事前评估对政策制定者的启示，包括绿色预算及其他方面。第5.2节将进一步展望统计物理学框架的扩展及其在其他政策情境中的应用。表A.1 概述了本文使用的跨学科术语；图6 和7 则分别展示了玻尔兹曼框架的整体流程图，以及在其有效描述中进行项目选择的示意图。

2 历史数据、项目概况与绿色分类

以下数据准备旨在指定作为统计分析实证基础的历史项目库。本节所作的选择将影响事前评估，因此必须以非常透明的方式记录。玻尔兹曼框架并不试图消除这一阶段所采取的“分岔路径”所带来的模糊性 [16]。它既不会优化指定的项目库，也不会以其他方式对其进行修正，而是将其视为毋庸置疑的输入。

为进行透明的示范，我们使用世界银行基础设施私人参与（PPI）数据库 [17]。这一选择是基于该数据库包含大量可公开获取、开放访问的数据，而非出于对世界银行或其机构及政策优先事项的任何特别偏好。虽然原始数据库查询的时间跨度最早可追溯至二十世纪初，但我们的分析范围限定为一个四十年的观测窗口，即1984年至2024年。这一时间截点并非随意选定，因为它标志着全球发展项目管线中有效且可按结构分类的绿色基础设施项目在历史上的起点。查询中的所有项目均涉及发电领域 [18]。

2.1 从原始数据到可观测量和项目概况

PPI数据库中的基本实体是彼此独立的项目，当然，其中存在一些冗余，我们将在下文予以筛除。在宏观描述中，相较于经过加工的行政或计量经济学指标，我们优先采用能够在单个项目层面得到严格记录的属性。除了可观测性这一要求（即数据中的每个项目都具有确定的属性值）之外，我们对将要分析的属性不施加其他限制。尤其是，后续分析并不要求我们假设随机变量或其底层联合分布的存在。

我们将此类直接观测到并选作分析对象的属性称为可观测量。确定一组可观测量（如X、Y和Z）后，每个项目都由相应的$(x、y、z)$元组表示，从而唯一确定其相关项目概况。在本文中，我们始终用大写拉丁字母表示可观测量，用小写元组表示给定数据集中记录的相应项目概况。

显然，若将某个没有信息价值的行政标记（例如 `Project ID`）用作可观测量，那么以此为条件就能区分每个项目，从而无法对集体行为作出任何有意义的贡献。通过将此类用于识别项目的标签排除在政策问题之外，具有共同特征的项目概况便会被归为一组，其经验权重和反事实权重分别由经验分布和假设分布体现，详见第 3 节。

物理可观测量。每个项目概况都包含项目铭牌容量的数值 y `Physical Capacity` Y，单位为兆瓦（MW）。为确保示例清晰明了，我们尽可能少地预处理记录值。对于 y 条目没有信息价值的项目，例如 `NaN` 或 `N/A`，我们将其排除，而不进行插补。物理容量Y 直接取自融资交割时记录的铭牌容量，不针对随时间变化的发电量、季节效应、天气条件或其他环境因素作进一步调整。此类改进可在后续阶段单独引入，而不会改变模型无关分析的核心。

在制定政策任务的过程中，我们保留物理可观测量的原始单位，而不将其替换为人为构造或推导得出的度量型指标。如果政策任务要求提供多个可观测量，例如铭牌容量与额定频率或储能容量，则只需将所有相应的可观测量保留

在项目概况的定义中，分别以兆瓦、赫兹或兆瓦时记录其数值。同样，也可以在铭牌容量之外补充数量信息（例如发电机组的数量），或补充燃料或能源资源的类别型说明，而无需进行数值编码。

通过复合表示对异质可观测量进行编码，可能会引入偶然依赖，从而以不可控的方式扭曲后续统计分析。幸运的是，玻尔兹曼框架能够在同一形式体系内处理由任意有限数量的可观测量定义的项目剖面。无需通过额外的建模选择将可观测量归约为共同的检验统计量，内在的 p 值（第 3.3 节）可直接通过对包含这些项目剖面的备选库的抽样概率进行粗粒化来构造。这样，无论定义项目剖面的可观测量数量或类型如何，玻尔兹曼框架都能保持其统计分析及后续评估的形式不变。

经济可观测量。由于我们的数据跨度超过四十年，因此自然存在一个时间可观测量 T ，用于表示项目年份 t；我们直接从记录的 `Investment Year` 中提取该量。尽管 T 体现了项目记录的时间方向，但它主要是一种经济可观测量，因为它记录的是投资年份，而非项目的物理年龄。

与绿色预算编制的财政政策维度相关的主要经济观测量是 `Total Investment` X，以百万美元（美元）计量。由于不同年份记录的货币价值以名义价值表示，我们通过将每项投资x 按年份$t \in [1984, 2024]$换算为2024 年不变价百万美元，以考虑观测窗口内的通货膨胀：

$$x_i \quad \to \quad c(t_i) \cdot x_i \qquad \forall\, i \in \{1, \ldots, N\}\,, \tag{2.1}$$

其中，时间调整因子 $c(t) \in [1, \infty)$ 定义为消费者价格指数（消费者价格指数）之比：

$$c(t) = \frac{\mathrm{CPI}(2024)}{\mathrm{CPI}(t)}\,. \tag{2.2}$$

请注意，这种重新缩放直接遵循金融和经济分析中对不同年份货币价值的标准处理方式。实际上，这种及类似的转换可以直接纳入控制财政方面的条件中，或通过下游指标实现，同时保留原始数据不变。除这一显而易见的调整外，我们不会对用于构建项目概况的记录值进行进一步转换。

2.2 绿色分类和政策可观测量

政策分类体系的实施——例如法国的五级尺度，或欧盟委员会将能源资产划分为可持续（绿色）、有害（棕色）或过渡性（中性）三个不同类别的框架——通常是一项广泛的

初步分析。这一初始过程通常通过技术专家、工程研究人员、机构利益相关者与政策从业者之间的审议，同时纳入科学考量与监管考量。

我们的统计框架建立在这一审议之后。它将已采纳的绿色分类法作为指定输入，并将其编码为政策可观测量 S，而不对该分类法背后的政治、监管或科学标准作任何判断。与学者中常见的一种观点相反，指定数据并不存在不确定性。如果某项被归类为绿色的技术后来被重新归类为环境有害，那么最初的分析就应视为错误，并须根据更新后的分类重新评估。

玻尔兹曼统计框架能够容纳 S 的不同形式。颜色分类方案意味着 S 是类别型的，而将项目容量中被标记为绿色的比例作为度量可观测量。为使这项初步研究清晰易懂，我们选择采用一种基于项目 `Technology` 的二元标记系统：根据每个项目档案中记录的技术字符串ξ，我们判断其中列出的发电技术能否明确且一致地归入环境友好/有害分界的任一侧，且不受多燃料技术带来的歧义影响。随后，我们根据以下规则为档案赋予二元政策指标 s

$$s = \texttt{green renewables} \text{ 如果} \xi \in \Big\{ \text{太阳能光伏；太阳能聚光热发电；聚光光伏；风能；地热；小型水电（<50兆瓦）；大型水电（>50兆瓦）；沼气；生物质；废弃物；太阳能光伏和太阳能光伏；太阳能光伏和沼气；太阳能光伏和风能；太阳能光伏和不适用；风能和太阳能光伏；风能和不适用 \Big\} \tag{2.3}$$

and

$$s = \texttt{brown/grey fossil} \text{ 如果} \xi \in \Big\{ \text{柴油；煤炭；天然气；蒸汽；柴油和煤炭；柴油和天然气；天然气和柴油；天然气和蒸汽；煤炭和柴油；煤炭和天然气；煤炭、柴油和天然气；不适用、柴油和不适用} \Big\} \text{。} \tag{2.4}$$

这种简化的划分仅根据低碳发电技术，将基础设施项目与传统的高碳资产基准区分开来。

为便于在后续分析中对绿色分类进行细粒度控制，我们将 `Technology` 属性中的原始文本映射为辅助政策可观测量S。这一额外的可观测量用于区分被统一标记为绿色的具体绿色技术。对于每个项目概况，值 $\tilde{s}$ 由相似的绿色技术分组确定

技术类型如下：

$$\tilde{s} = \begin{cases} \texttt{not applicable (N/A)} & \text{if } \ s = \texttt{brown/grey fossil} \\ \texttt{hydroelectric} & \text{水电} \in \xi \\ \texttt{solar fleet (PV/CPV/CSP)} & \text{太阳能} \in \xi \\ \texttt{onshore wind} & \text{风能} \in \xi \\ \texttt{geothermal heat} & \xi = \text{地热} \\ \texttt{bio-energy \& waste (biomass/biogas/waste)} & \text{else} \end{cases}, \tag{2.5}$$

其中， ξ 表示记录在 Technology 下的原始字符串条目，s 表示上一段中定义的绿色标签。

至此，我们强调，玻尔兹曼统计框架同样可以直接作用于未压缩的 Technology 属性，并据此进行边缘化，而无需引入主要政策可观测量和辅助政策观测量。此处将主˜要政策可观测量 S 与细粒度政策可观测量S 加以区分，只是为了便于下文呈现和讨论政策情景。显然，这两个可观测量并非彼此独立的信息来源，因为二者均由同一 Technology 属性确定性导出。

事实上，人们同样可以定义单一政策可观测量，使其包含五个归组的绿色技术类别以及一个汇总的棕色类别。在该描述中，统称为绿色标签的条件样本均值在表述强制条件时，必须显式地对五个绿色类别进行汇总。双可观测量描述将绿棕分类与绿色技术更精细的构成区分开来，从而能够以主要政策可观测量清晰表述强制条件，同时保留细粒度政策可观测量以供下游分析。

最后，为确保绿色分类明确无歧义，需要严格的排除过滤器，以防止含糊归类并消除多燃料混杂。包含争议性技术（Nuclear）或绿色-棕色混合组合（例如Hydro， Large (>50兆瓦) & 柴油或 Natural Gas & Steam & Solar-PV）的项目，连同 ξ 包含无信息元数据（Not Applicable、 Other以及缺失的 NaN 变体）的项目，均被系统地排除在样本之外。

2.3 次级可观测量与完整项目概况

˜除核心T， X， Y 以及政策可观测量S， S （上文已引入）之外，本研究还考察了以Z_i 表示的进一步可观测量，对应于 Investment Share、 Region以及IDA Status 等世界银行数据库属性。这些可观测量既不直接进入绿色强制要求的表述，也不直接进入第3.1节中的结构条件。尽管如此，所有次级可观测量的取值均构成完整项目概况的一部分，因此与主要观测值一样，须经过同样严格的数据清洗（不含无信息或损坏的取值）。

供后文参考，本研究选取以下可观测量来定义一个完整的

项目概况可分为三类。首先，核心可观测量包括物理可观测量Y，表示铭牌容量，以及财政经济可观测量X 和T，分别表示投资额和年份。其次，我们定义˜政策可观测量S 和S，分别表示绿色—棕色标记和细粒度的绿色技术类别。最后，次级可观测量Z_i 表示与经济和地理属性相关的其他特征，用于开展事前分析，以评估绿色强制要求的下游影响。

在反事实项目库中，所选次级可观测量的统计数据，与核心可观测量和政策可观测量的统计数据共同确定，且满足绿色强制要求及任何指定的外生约束。随后可在下游分析中考察由此得到的联合统计数据。此处引入上述分组，仅为便于记录和阐述。它既不意味着，也不要求我们设想这些可观测量的有效描述或其联合行为的任何结构化模型存在相应的分解。对于玻尔兹曼构造而言，重要的是定义项目概况的可观测量集合，以及能否根据其中部分或全部可观测量，以定量方式表述政策假设或结构约束。某个可观测量不必纳入任何一个（或每一个）假设；将其归为核心、政策或次级，也不意味着它在反事实构造中具有任何方法论上的不同作用。

2.4 历史项目库

本节最后给出基础设施项目历史库的最终形式。一旦达成一致，分析过程中就不得更改它；也就是说，实际上不会再有归因于这一历史基线的剩余不确定性。

识别不同项目。处理跨国项目数据库时，一个经验性挑战是存在重复合同记录、拆分的融资分期以及反复出现的交易记录。若不加以校正，这些条目可能导致物理容量被重复计算，并使记录的投资额虚增。具体而言，尽管铭牌容量相同，同一资产仍可能以多个数据库条目的形式出现在连续的投资年份中，且资本金额各不相同。在对数据库所有条目中的 `Project name` 属性文本进行标准化后，我们使用这一标签来识别不同项目：每当同一项目在数据集中出现不止一个条目时，我们首先核实各条目在关键不变可观测量上的观测值是否一致，例如 `Technology`、`Country`、 `ContractPeriod`等。随后，我们将这些条目合并为一条唯一的项目记录。该记录保留不变的观测值，同时汇总各笔资本配置，以得到总投资x，并选取最早记录的年份作为该项目记录的年份t。

最终，这一识别过程减少了数据集中的条目数。更重要的是，识别不同项目不仅是为了避免物理容量被重复计数，还决定了历史存储库中抽样单位（项目）的实际数量N 。如果多个数据库条目对应同一个底层项目，却将它们视为彼此独立的观测值，就会人为地增加N，从而在引入不合理的——

敏感性并影响项目概况的经验频率。该N 还设定了数据集中可观测的不同项目概况数量的上限。

由此得到的样本与接受阈值。 在应用前述去重和筛选规程，且不引入任意数据填补或合成变量后，我们得到一个数据集，它最大限度地保留了自定义查询的原始经验统计数据，未作进一步修改。由此得到的抽样规模为

$$N = 3\,729 \text{ unique infrastructure data points} \tag{2.6}$$

这为采取保守的接受阈值（即传统假设检验中所谓的显著性水平）提供了依据

$$\alpha = \frac{1}{N} = 0.000268 \approx 0.03\% \ll 5\% \,. \tag{2.7}$$

该阈值为第 4.1 节中反事实情景的假设检验设定了严格的决策边界，因此，统计拒绝所需的证据必须显著超出与单个抽样单位（项目）相关的水平。

为避免常见的误解，我们强调，我们不会对连续的度量可观测量进行分箱、平滑或其他处理，例如物理容量Y、资本配置X，也不会处理项目年份T等时间观测量。对`Technology` 属性重新分类，仅用于通过二元政策可观测量S 及其细粒度对应项 $\tilde{S}$来编码所采用的绿色分类。这种分类属于外生的政策选择，并非统计框架中的任何数学约束所决定；该框架也能够处理度量型政策可观测量。分配给S 的类别也可以依据相关政策过程所商定的任何其他分类体系来定义，而不改变所实施的公式。

3 玻尔兹曼抽样的反事实世界

有了经过透明处理、用于指定历史存储库的完整数据集，我们现在转向构建反事实政策情景。对政策制定者而言，这类反事实提供了理想化的政策全集，可用于考察某项政策目标可能带来的后果，而无需在基础设施中实际实施相应干预。

为利用第 N 节所概述的预期大1.3简化，我们采用 [15, 19]中提出的组合框架。我们并不预先假定抽象的概率空间或潜在总体分布，而是从观测项目的有限集合出发，构建一个反事实库。在政策语境下，我们将一个能够完全实现既定政策目标的反事实库称为假设反事实。在这个理想化的项目库中，可将其视为一个无限大的瓮，项目本身充当抽样单位（或瓮中的球），并且可根据其

剖面（即球的类型）进行区分；这些剖面由针对政策问题和预期下游分析所选定的可观测量定义。不同的假设反事实对应于不同的政策宇宙，在这些政策宇宙中，项目选择的未知机制可以用一般抽样进行有效描述。

在本研究中，完整的项目概况由核心、政策及次˜级可观测量Y、T、X、S、S 和Z_1, Z_2, Z_3 分别决定，如第2节所选。这样一来，每一种不同的剖面（y、t、x、s、 ˜s、z_1, z_2, z_3）都标识瓮中的一种球。假设分布 $\mathfrak{p}$ 为原则上可由该八元组索引的每一个可观测剖面赋予一个权重，从而仅通过重加权来规定假设反事实的构成。在大N 情形下，从该假设反事实中抽取的、由N 个项目组成的替代项目库，至少就预期影响分析的目的而言，可被视为一般抽样的结果，其中各项目系独立抽取。

非公理框架并不要求事先将可观测量声明为度量型或名义型。某一可观测量进入假设时所采用的模态，在逻辑上取决于施加于其上的条件：关于均值或方差的条件自然赋予其度量模态，而边际流行度则赋予其分类模态。反之，一个条件也可能涉及度量可观测量的特定取值或类别。在玻尔兹曼构造中，这些区分并不影响有效的重加权机制，该机制对这些区分并不敏感。因此，可观测量的模态问题，就归结为如何将具体的政策授权表述为涉及该可观测量的条件。

我们的统计物理学视角还有另外两个特征，对于当前这项对底层项目选择过程仅有极少知识的应用尤为重要。在一般抽样的普适类内对抽样概率进行粗粒化，可给出一个内在的p值，而无需假定总体模型，也无需引入特定于应用的检验统计量并渐近地确立其参照分布。在推导出能够在大样本量N 下稳健计算该p值的半解析闭合公式之后，我们不再需要依赖计算密集的蛮力自助法来生成合成数据集。以统一的方式，为每个反事实政策情景构建的假设分布既使其下游分析成为可能，也直接进入决定该内在p值的检验统计量。

经验立场。经验分布记为 $\mathfrak{f}$，将相对频率赋予第2.4节所指定的历史存储库中观测到的各个不同项目概况。这些概况定义了我们思想实验之瓮中的球类型。请注意，项目描述属性中的“基本”成分，即其概况，完全取决于当前的评估。不同的政策任务和下游指标通常要求不同的可观测量，因而要求不同的项目概况，正如第2.3节所述。

除非历史存储库中观测到了所选可观测量各自取值的所有可能组合，否则观测到的项目概况集合就无法整齐地分解为各个可观测量的值集。换言之，观测到的项目概况集合无法穷尽各个值集的笛卡尔积，更遑论在这些数学域之间呈现某种积结构。这种情况的发生

是由于有限样本量，或是由于分析者可能并不知晓的结构约束，导致某些组合在技术上或行政管理上不被允许。例如，预算合理的模块化绿色电力设施不可能拥有大型水电站的铭牌容量。

由于决定不向未见过的剖面进行外推（这需要更深入的理解），经验分布 $\mathfrak{f}$ 具有严格为正的权重。更值得注意的是，剖面集合的非乘积结构使得难以根据规定的一般化矩推导出假设分布的闭式表达式，即便在线性条件较为简单的情况下也是如此。在以模型为中心的方法（另见最大熵方法）中，这乍看之下似乎是不理想的特性。然而，正如我们在第 3.2 节所论证的，恰恰是这种非平凡的经验结构，才为事前政策问题带来了明确的答案，并能轻易确保可归因性。在实践中，无法根据可控参数写出假设分布的闭式表达式，并不影响基于 I- 投影（第 3.2 节）构造假设反事实，也不影响利用 N 渐近半解析闭式表达式对其进行统计评估，该表达式用于计算 p 值（第3.3 节）。

3.1 绿色强制要求假设

尽管我们的框架可以轻松纳入涉及多个目标可观测量的更复杂政策任务$Y_1, Y_2, \ldots\ldots$，但此处我们有意采用铭牌容量这一较简单的基准示例Y 见第2.1节2.2节。下文中，我们将政策任务设定为：第2.2节中标记为绿色的总铭牌容量超过标记为棕色的铭牌容量。

具体而言，我们评估执行电力平衡 $\eta = 0$ 的影响，并考察绿色电力盈余 $\eta > 0$ 或赤字 $\eta < 0$（用于回顾性一致性）的不同目标水平；这些目标由预期电力差表示，

$$\eta \stackrel{!}{=} \sum_y y \cdot p(y, \texttt{green renewables}) - \sum_y y \cdot p(y, \texttt{brown/grey fossil}) \,. \tag{3.1}$$

此处，假设分布 $\mathfrak{p}$ 描述了反事实库的构成，为每个观测到的项目概况指定一个特定概率（即非负权重）：$p(y,\ t,\ x,\ s,\ \tilde{s},\ z_1, z_2, z_3)$；每个观测到的项目概况为$(y,\ t,\ x,\ s,\ \tilde{s},\ z_1, z_2, z_3)$。为简化记号，上文中诸如$p(y,\ s)$ 这样的简写索引仅使用部分可观测量，它被理解为 $\mathfrak{p}$ 的相应边缘分布，该分布通过对其余可观测量所有联合观测值求和得到。

使用经验分布 $\mathfrak{p} = \mathfrak{f}$ ——直观地说，就是我们指定的历史存储库的多维直方图，我们可以直接得出式（3.1）中的样本均值之差，

$$\mu = \sum_y y \cdot f(y, \texttt{green renewables}) - \sum_y y \cdot f(y, \texttt{brown/grey fossil}) \,. \tag{3.2}$$

历史的　与负差值相关的绿色电力缺口无法符合　h

旨在实现绿色电力盈余（即功率差为正）的前瞻性政策目标。即使历史基线中的差值为正，也不意味着自动符合强制目标。符号或幅度上的这种不匹配，促使我们构建反事实资料库。反之，如果历史基线已经达到强制目标，就没有必要构建某种“无偏”的历史存储库版本，因为历史存储库仍然是唯一指定的事实基准。正如我们将在第3.3节看到的，内在的 p 值在这种情况下将变为1。

结构条件。在我们的示范情景中，为反映现实世界中的运营和制度约束，我们需要保持特定的总体物理和财务量不变。在玻尔兹曼框架中，这些外生约束被表述为对可接受的假设反事实的结构条件。例如，监管机构在评估政策转变时，可能会锁定总资本预算或某项技术的专项资金分配。

作为一个直观范例，我们评估向绿色电力盈余转型的过程，同时保持所有区域和项目批次的总铭牌容量不变。从数学上说，我们固定可观测量Y的一阶矩，因为N 乘以其样本均值，决定了历史存储库中的总容量。理论上，这个一阶矩可以校准为观测到的概况所支持的任何物理上合理的数值；但由于缺少先前的行政指令，我们将其设为经验值：

$$\sum_{y} y \cdot p(y) \stackrel{!}{=} \sum_{y} y \cdot f(y) \,. \tag{3.3}$$

对Y 一阶矩的全局条件，可以理解为对新增发电容量交付的一项基本物理约束。在当前情境中，Y代表历来通过世界银行项目承诺的新增铭牌容量。维持其总体历史水平，可确保政策授权本身不会导致新增铭牌容量相对于历史承诺增加或减少。从工程视角来看，视电力系统中已有的容量而定，任何一个方向的偏离都可能分别造成发电限电或供应短缺。

显然，要对全球电力系统的实际情况进行详细分析，还需要有关其物理结构和运行的额外信息，包括区域需求、调度、储能、输电与互联容量，以及区域间的电力流。此类分析超出了本示例的范围，单靠此处使用的历史存储库无法完成。特别是，在具备相应信息且政策要求明确时，可将区域容量约束纳入额外的结构条件。

由于假设反事实是通过一个规定某个不一定有限的项目库构成的概率分布来定义的，因此很难将式（3.1）中的重新缩放的均值差解释为总体绿电与褐电功率差。当总容量固定后，相应的功率差占总容量的份额则提供了

一个定义明确的无量纲度量，用于表示强制性绿色电力盈余或赤字。该归一化量在第 4 节的各图中始终作为控制参数使用。因此，在总容量固定的情况下，绿电与褐电功率差和绿电与褐电功率比（只要后者有明确定义）就成为表示绿色强制要求的等价参数化方式。具体而言，该比值可以完全用式（3.1）中的目标功率差以及式（3.3）中的固定总容量表示。

此外，我们在反事实中引入一项运行边界条件。借助细粒度政策可观测量 S，我们通过严格不变矩，在整个绿色转型过程中保持水电容量不变

$$\sum_{y} y \cdot p(y, \texttt{hydroelectric}) \stackrel{!}{=} \sum_{y} y \cdot f(y, \texttt{hydroelectric}) \,. \tag{3.4}$$

水力发电的物理、地理和环境特征构成了这一外生约束的依据。与模块化可再生能源技术不同，水电容量无法根据政策授权任意扩张。一个国家不能仅凭工程手段改造高海拔河流或湖泊来修建大型水坝。事实上，加速的气候变化和频繁发生的干旱直接威胁着现有水库，进而也威胁着水电基线。鉴于这些问题，我们不允许反事实中的绿色转型依赖于超过其本已脆弱的历史水平的总水电承诺。

组合假设。式（3.1）中的强制目标与式（3.3）和式（3.4）中的结构条件共同定义了政策目标的组合假设。由此，组合假设确定了观测到的项目概况上的一个分布线性族，其中的分布满足绿色强制要求和所有施加的结构条件。该线性假设族中的任何分布都明确描述了一个满足组合假设的假设反事实。

在大N 情形下，借助世界银行等国际组织维护的大型项目库，这一情形是可以实现的，不同项目概况的数量也可以随N增长。这种同步扩展反映了大多数政策问题的多维特性，因为这些问题旨在分析后续的多个可观测量。反之，一个只包含例如三个项目、且受两个独立条件约束的历史存储库，规模太小，根本不值得进行统计评估。在这种情况下，若不引入额外假设，几乎无法保证可归因性。

大多数政策问题只围绕少得多的维度来制定，因而转化为少数几个相互独立的强制性条件和结构条件，这些条件共同定义了组合假设。（在我们的示范性应用中，我们陈述了四个条件：一个强制性条件，以及在单纯形上进行归一化之外的两个非平凡结构条件。）这正是我们所关注的<style id='1'>N 较大情形，其具有可处理性的统计结构在文献中一直被完全忽视。无论项目概况的数量是否同时增加，并在极端情况下可能达到<style id='3'>N 个不同的项目概况，从而标志着经典 Sanov理论的失效，粗粒化抽样概率仍然可以表示为半解析闭式表达式 [15]。

定义欠指定线性系统的组合假设，在假设反事实的构成方面会留下剩余的自由度。当该线性系统至少允许一个分布时，就称该组合假设在历史项目概况的范围内是可行的 。这表明，无需调用超出历史基线的未观测项目概况，即可实现政策目标。在欠指定情形下，可行的联合假设的线性族包含无穷多个分布。

3.2 玻尔兹曼假设分布

组合假设规定了假设反事实必须满足的条件，但在常见的欠充分设定下，并不能确定唯一的假设反事实。这种简并性表明，仅由强制政策推导出的条件，即使辅以外生约束，也不会自动导出假设反事实，因此可能产生多个下游答案。

值得庆幸的是，为了使内在的 p 值具有可解释性，我们将经验分布 $\mathfrak{f}$ 投影到由组合假设定义的线性族上，采用前向 I-投影所得的分布作为假设分布。只要假设线性族非空，该投影就是唯一的。因此，对于每个可行指令，我们都提供了一项有原则的选择准则，用以确定唯一的假设分布，即所谓的 玻尔兹曼假设分布，进而确定假设反事实的构成。具体而言，在下游分析中，这一玻尔兹曼假设反事实 被用作最小扰动情景；组合假设未明确规定的维度，则由 I-投影无歧义地确定。

在以下段落中，我们将阐述选择假设分布的核心认识论论证。在通过公共治理中的工作示例概述内在p值的含义之后，我们将陈述其在统计假设检验中的可解释性条件，并说明该条件如何在较大的N 下自然导向玻尔兹曼假设分布。最终，我们将展示，在玻尔兹曼假设反事实下，以历史绿色电力缺口为条件进行抽样，是由历史存储库的额外经验结构所决定的。这种条件抽样中额外结构的经验可归因性，是一种能够唯一标识玻尔兹曼假设反事实的属性。

一般抽样。 在大N 区间内，对于本次事前分析，可以通过一般抽样来描述促成项目库形成的多种机制的集体影响：从假设反事实中独立抽取项目。用统计物理学的术语来说，当系统的细节，或更一般地说，某个过程的细节，在可接受的近似范围内对所研究量的描述变得无关紧要时，这种情形称为普适性。普适性并非外部选定模型所赋予的属性，而是在相关极限下，现实应用本身所具有的特征。一般抽样通过多项方案提供了一种最简描述，无需指定促成项目库形成的各个具体机制。

拓展玻尔兹曼最初的思想实验，可以对一般抽样过程作出具体的物理实现。具体而言，我们考虑与历史存储库规模相同的备择存储库 N，这些存储库从假设反事实中抽取，并满足结构条件式（3.3）和式（3.4）。对备择存储库进行这种条件抽样，可以通过从相应的假设分布中进行拒绝抽样来实现。这些有限的项目库代表在选定的假设反事实下生成的替代历史版本。

由于抽样不确定性——在瓮模型中体现为现实世界影响因素的集体效应——总功率差方程（3.1）会因备择假设的项目库而异。因此，根据这一总功率差，将备择假设的项目库划分为不同的存储库族是自然而然的。按照构造，属于同一族的所有备择假设项目库都支持方程（η 中的同一强制要求（3.1），并且在方程（3.3）和（3.4）中具有相同的广义矩经验估计。其中一个族呈现出历史绿色电力缺口。根据定义，历史项目本身属于历史存储库族。

信息投影的出现。通过在条件抽样方案中对备择假设项目库的概率进行粗粒化，每个存储库族都会被赋予一个概率质量，该概率质量取决于所选择的假设反事实。若适用大N 条件，则该概率质量的鞍点近似决定事前评估的主导阶结论，而对项目库形成进行更详细描述所产生的影响则出现在次主导阶。内在p值是这样一个累积概率：从假设反事实中抽样得到的任意存储库族，其发生概率至少与观测到方程（3.2）中绿色电力缺口所对应的历史存储库族一样低。

每当N 相对于假设条件的数量足够大时，便可从这一内在p值普遍得出结论；该值允许进行 χ^2 近似。正如我们在 [15], 中所推导的，由此得到的检验统计量等于N 的两倍乘以I投影。在当前这一涉及单一强制目标的应用中，内在p值借助自由度为1的 χ^2 分布的累积分布函数加以近似。I投影是通过在历史存储库族上最小化相对于所选假设分布的信息散度而得到的，因为我们已将分析限定于能够再现结构矩经验估计的假设分布。

认识论要求。 在不依赖外部模型假设的情况下，我们已为这一原本高度复杂的现实问题确立了这一可处理性水平，接下来必须确定进入p值计算的假设分布。为此，我们考察内在p值在事前政策目标统计评估中的操作运用。随后的认识论考量显然适用于任何使用基于抽样的p值所进行的统计评估，而与具体应用无关。

要使内在的p值能够依据预设的接受阈值支持有意义的决策，历史存储库必须是在所有能够解释观测到的绿色电力缺口的备择假设中，概率最高的那个存储库——

。否则，较高的p值可能会矛盾地暗示，观测到的缺口在假设反事实下具有足够高的概率，尽管据以得到该缺口的历史存储库本身看起来却极不可能，尤其是在N较大时。必须避免这种矛盾的解读，因为指定的历史存储库毕竟是事前评估中唯一不变的要素。

我们将这一要求视为解释内在p 值的认识论准则。当 N 较大时，只要要求历史项目库是历史存储库族中概率最大的成员，就可以排除上述悖论情形。如附录 A.2 所示，最大概率这一要求由玻尔兹曼假设分布满足：该分布是经验分布 $\mathfrak{f}$ 在由组合假设定义的线性族上的正向 I 投影。只要组合假设可行，相关的 [20], I 投影就是唯一的，从而提供了一种有原则的方法来选择假设反事实。

如果我们纳入关于项目库形成过程的更详细信息，从而采用不同的抽样方案，例如在假设反事实中考虑有限总体效应的广义波利亚 urn，那么在物理尺度的适当极限下，就会出现另一种散度度量（见 [15] 的附录 A）。归根结底，这会在相同的认识论准则下确定另一种假设分布。若要进行可归因的事前影响评估，鉴于结果取决于所采用的抽样方案，因此必须透明地说明纳入了哪些资料库形成细节，因为不存在任何认识论原则能够规定这些细节。

一幅辅助图景。在当前广义矩的线性条件以及大N下有效的多项抽样设定中，认识论准则选出的玻尔兹曼假设分布，也可以通过条件多项抽样方案来理解。我们反复掷一枚不公平的多面骰子，其各面的概率由观测项目剖面的历史相对频率给出；只有当掷出的结果满足组合假设时，才记录由此产生的无限大反事实项目库。在这些可接受的反事实中，组成由玻尔兹曼假设分布决定的项目库，是理想化无限大小极限下唯一以非零概率存续的项目库。这一辅助图景源自对条件抽样概率的经典萨诺夫型展开，其中经验分布 $\mathfrak{f}$ 如今作为参照。广义矩的非线性条件或其他抽样方案未必能以抽样为基础来理解认识论准则所产生的后果。

反事实广义矩对历史基线的可归因性。在实际应用中，项目库中还有许多维度未被第 3.1 节的组合假设直接指定。这些额外维度可由广义矩集合来表示。假设反事实会为它们指定具体取值，而这些取值进而会影响下游指标以及内在的 p 值。除非

要求在历史存储库中得到完全满足，否则反事实取值将不同于历史基线。

在各种可能的假设反事实中，只有其组成由玻尔兹曼假设分布决定的那一种，才会为所有额外的广义矩规定取值；在条件抽样下，这些取值完全可归因于相应的历史估计值。这一次，条件方案涉及对嵌套于历史存储库族中的备选存储库族的抽样概率进行粗粒化。只有在玻尔兹曼假设反事实下，给定历史绿色电力缺口，抽取历史估计值或任意额外广义矩中至少同样不可能的取值的内禀p值才等于一。这类嵌套评估不应与对历史绿色电力缺口本身的统计评估混为一谈；后者是在结构条件下评估该缺口。

具体而言，我们将嵌套仓库族定义为：除满足结构条件方程（3。3）、（3。4）以及历史绿色电力缺口条件方程（3。2）外，还将进一步的广义矩约束为其历史估计值。在同一假设反事实下，我们相对于外围的历史存储库族，对该嵌套仓库族进行统计评估。可归因于历史基准意味着，这项统计评估的内禀p值必须始终为一（在鞍点近似下），无论采用哪一个额外的广义矩。

上文所引入的内在p值可解释性的认识论论证，也可被理解为对历史存储库之组合附加结构的可归因性的极端情形。为此，我们指出，历史存储库是该嵌套族的唯一成员；该嵌套族通过指定其余维度的历史估计值来定义，这些维度在观测到的项目概况这一有限集上补充结构条件与任务条件。对于这一被最大限度指定的嵌套族，其p值在玻尔兹曼假设反事实下自动为1；在该反事实中，历史存储库在所处的环境历史存储库族内具有最高的抽样概率。

线性代数形式与数值计算。由于观测到的项目概况数量以N为上界，分布 `p` 可由包含不同剖面概率的列向量表示。相应地，形如式（3.1）那样、用于规定由 `p` 所描述的反事实中样本均值的条件，可通过此类概率列向量与一行向量的标量积来实现，该行向量的系数或者等于y（对于 `green renewables`），或者等于 $-y$（对于 `brown/grey fossil` 剖面）。例如，为计算 `hydro-electric` 技术的普及程度，我们会改用一个行向量：对于具有 $\tilde{s}$ = `hydro-electric` 的项目概况取1，否则取0。归一化条件由对所有项目概况均为1的常数行向量给出。

为概括一组条件，我们可以使用由上述行向量构成行的系数矩阵。将该系数矩阵作用于概率列向量，会得到另一个列向量，其中包含相应广义矩的值。上述线性代数形式 [19] 为组织条件信息散度最小化提供了一种便捷方式。具体而言，在

剔除冗余条件后，我们可以稳健地确定参与后续假设检验的相关自由度。在当前情形下，这些自由度等于组合假设的系数矩阵的行秩减去包含结构条件的系数矩阵的行秩。

在不存在因组合假设而权重被强制为零的项目概况时，玻尔兹曼假设分布属于由系数矩阵生成的指数族，并以 f为参照。这种熟悉的数学形式并非强加于反事实的外部假设。相反，它源于坚持为我们的统计评估采用可一致解释的内在p值。顺便指出，由此得到的相对于原始数据的最小偏差选择恢复了Efron 的指数倾斜 [21] 在自助法设定中的结果，尽管无需下游自助法。

利用其指数形式，我们可以在滤除线性相关条件后，依据独立条件对玻尔兹曼假设分布进行参数化。相应的剖面概率随后可借助牛顿-拉夫逊求根算法进行数值确定。[2] 这种线性代数参数化通过一种快速而稳健的数值程序简化了前向I投影的估计，该程序用于确定线性条件下信息散度的非平凡最小值。

3.3 运营含义

对于特定的绿色电力盈余，原则上存在许多由观测到的项目概况构成的假设反事实，它们都与指定的结构条件相容。这些假设反事实在下游指标方面存在差异，例如投资分配或资本效率系数。这种答案的多样性使得有必要建立一个独特的参照点，以供民主协商和决策使用。

拥有玻尔兹曼假设分布是所提出范式转变的核心收益。在不具体规定项目实际选择方式的情况下，玻尔兹曼假设反事实的构成依照其构造，可归因于历史基线以及由结构条件和强制目标构成的组合假设。尤其是，这种可归因性确保了附加性的构成层面。在公共政策评估中，附加性通常与政策归因联系起来讨论 [22, 23]。在玻尔兹曼框架中，不存在政策目标意味着不会偏离历史存储库；另请参阅第3节引言部分关于经验立场的段落。不过，在规定的设定下，反事实对历史存储库的偏离完全由强制目标触发，这并不意味着或断言关于没有政策目标时将会发生什么的因果反事实。

理想化政策全集中的替代历史版本。 玻尔兹曼假设反事实描述了政策制定者设想的政策全集。在这个与历史存储库仅有最小偏离的反事实项目库中，项目构

[2]如需优化 `C++` I-*projector*的实现，请参阅 `Python/R` 封装器，网址：
https://gitlab.uni-marburg.de/fb20/ag-chung/iprojector/。

成满足强制性绿色电力盈余，同时遵循任何指定的结构条件。（我们强调，假设反事实并不代表基于历史基线的预测。）在这个理想化的全集中，假如政策目标已获采纳，政策制定者会追问：已采纳的政策遇到影响现实项目库的诸多因素时，会发生什么。行政惯性、零散的国别政策、社会经济发展、政治决策、工程约束和气候波动，都会影响哪些项目得以落地，以及各自所占的比例。

由于这类情况发生的方式无法被完全预见，也无法详尽明确地描述，因此，反事实全集中可能出现、代表替代历史版本的有限项目库，不会与假设反事实的构成完全一致。例如，财政可观测量 X 的一阶矩代表给定备择项目库中的项目成本估算，通常会不同于假设反事实中的预期估算。尽管如此，对广义矩期望 [3]，而言，将其在各个备择项目库中的不同取值取平均后，便会在由玻尔兹曼假设分布给出的理想化政策全集中，重现其反事实估计。

沿用统计物理学对集体现象的视角，我们通过对与历史存储库规模相同的备择项目库 进行一般抽样，来表示尚未明确的影响所带来的不确定性。在瓮思想实验中，可以把项目想象成不同类型的球，各自对应不同的剖面。在这一图景中（图 7），从一个组成（各类型的比例）由玻尔兹曼假设分布给定的无限瓮中抽取 N 个球，即可形成一个备择项目库。

用于下游分析的综合框架。由于玻尔兹曼假设分布为所有观测到的项目概况赋予反事实权重，政策制定者可以计算任何能够用可观测量函数的期望值及其非线性函数表示的、数学上定义良好的度量型，包括此类期望值的非线性函数。第4.2阐释了这一下游评估的灵活范围及其显著的实际简化，这与以模型为中心的方法形成对比。在传统评估中，即使政策授权保持不变，考察不同的下游指标通常也需要建立新模型或扩展现有模型。在这里，假设反事实已经确定，使政策制定者能够将精力从决定还需要哪些额外模型，转向研究对其决策至关重要的后果。

公共治理中特别值得关注的是比例原则，该原则要求公共干预不得超出实现其目标所必需的限度。欧盟法律明确承认这一原则：《欧盟条约》第5条要求，欧盟行动的内容和形式不得超出实现条约目标所必需的限度 [24]。在监管影响评估中，会从有效性、成本和后果等方面审视各种政策选项 [2]。此类评估以具体的反事实

[3] 在对备择假设项目库进行条件抽样时，广义矩的期望值表示为对其可行值的求和，并根据由该矩的取值和结构条件定义的相应存储库族的粗粒度抽样概率进行加权。对于任意数量的历史项目概况，都可以通过低维高斯分布，利用大 N 鞍点技术近似计算该总和。

情景为前提，以便考察不同政策选择可能产生的后果。

在政策目标已明确的情况下，玻尔兹曼假设反事实提供了一个偏离最小的情景，可用于基于证据的比例原则事前评估。通过针对不同政策目标构建反事实，政策制定者可以定量比较这些目标所带来的影响。如果强制目标所隐含的总体最小偏离仍会造成被认为过度的后果，政策制定者就有明确的定量依据质疑这一强制要求本身。

统计假设检验。对于严谨的事前评估，内在的p值使政策制定者能够在指定的结构条件下，评估历史基线与强制政策目标之间的统计兼容性。不同于传统方法通过指定总体模型来确定所选检验统计量的抽样分布，我们的框架中的检验统计量源自项目存储库形成的大规模行为。 χ^2 近似并非脱离实际评估过程而独立设定，而是源于有效独立抽样的大N 情形。因此，我们p值的证据意义植根于其基于抽样的定义本身。在我们的示范性应用中，内在的p值量化了这样的累积概率：抽样得到的电力差异至少与玻尔兹曼假设反事实中的历史绿色电力缺口一样不可能。

在条件抽样方案中，我们仅保留构成满足指定结构条件、但除此之外由有限样本的组合方式决定的备择项目库。因此，不确定性并非来自共同认可的政策目标，也不是来自定义假设反事实的结构条件，而是源于有限样本存储库的构成围绕理想化项目构成发生波动。这些波动会导致总功率差偏离强制目标，同时仍满足结构约束条件。

在玻尔兹曼假设反事实下，内在p值汇总了历史存储库族的抽样概率，以及所有其他抽样概率不高于该历史族的可接受存储库族的抽样概率。因此，它不仅考虑呈现观测到的绿色电力缺口的备择项目库，也考虑属于抽样概率至少与历史存储库族同样低的存储库族的其他功率差。（在玻尔兹曼框架下，也可以构造以绿色电力缺口为目标的单侧检验。）

低于接受阈值的p值意味着：如果我们认可了界定玻尔兹曼假设反事实的绿色电力盈余，那么仅凭现实世界机制的共同作用，历史上的绿色电力亏缺极不可能出现。按照假设检验的标准逻辑，我们拒绝这样的假设：在历史数据涵盖的时期内，各方曾就这一绿色电力盈余达成一致。换言之，在所指定的结构条件下，历史基线在统计意义上与强制规定的绿色目标不相容（或不可信）。

正如美国统计学会所强调的， [25], p值并不能量化

某一假设为真或为假的概率，这意味着科学结论以及政策决策不应仅仅取决于其是否跨过预设阈值。遵循传统的费希尔逻辑进行假设检验时，[4] 我们的p值同样无法量化获得强制性绿色电力盈余的概率。相反，在特定绿色目标的组合假设下，它量化的是获得观测到的赤字或更不可能结果的累积概率。我们的贡献在于，在大N 状态下，根据现实世界机制对项目存储库形成的集体影响推导出这一概率，而无需引入特设检验统计量或外部指定的总体模型。

因此，在玻尔兹曼框架下，假设检验的结果可归因于政策目标，以及结构条件、历史存储库和缺乏有关项目选择的更具体信息。鉴于这种可归因性，若显著的p值（低于接受阈值），则应主动重构历史基线，以实现强制目标；前提是政策制定者在评估下游后果后决定实施该目标。

4 结果

鉴于已有历史存储库和政策目标，上一节提出的方法可直接用于构建和评估反事实政策情景。根据绿色目标和结构条件（共同构成绿色强制要求的组合假设），对观测到的项目概况进行重加权，可得出玻尔兹曼假设反事实。由此，可以针对历史存储库中已有的任何维度，评估强制要求带来的后果。对玻尔兹曼假设分布进行抽样，也可得到p值，用以量化强制政策下历史基线的合理性。

现在，我们将这一统计框架应用于世界银行项目库中的历史发电项目。我们探究与可行组合假设相对应的绿色电力盈余的完整范围。除了通过图表展示的标准财政和社会经济指标事前影响评估外，我们的下游分析还进一步评估了在规定的绿色电力盈余条件下，历史绿色电力缺口的合理性。

4.1 绿色转型的统计评估与财政影响

对于式（3.1）中的每个目标功率盈余或赤字，我们通过数值计算确定玻尔兹曼假设分布 $\mathfrak{h}$，其依据是

$$\mathfrak{h} = \underset{\mathfrak{p}\in\mathcal{H}}{\operatorname{argmin}}\, D(\mathfrak{p} \parallel \mathfrak{f}) \,, \tag{4.1}$$

[4] 有趣的是，p 值按其最传统的定义，竟然是我们唯一能够在拉普拉斯近似下推导出闭式表达式的粗粒化抽样概率。

其中 $\mathcal{H}$ 是由第3.1节中的组合假设定义的假设线性族，这意味着每个分布 $\mathfrak{p} \in \mathcal{H}$ 都满足式（3.1）、式（3.3）和式（3.4）。按照第3.2节概述的线性代数表述，这种条件最小化会为每个组合假设生成一个列向量，其中包含由 8 元组 $(y, t, x, s, \tilde{s}, z_1, z_2, z_3)$索引的非负权重。

由于数值最小化的计算成本很低，我们在式（3.1）中，对密集取样的一系列目标绿色与棕色电力差异计算玻尔兹曼假设分布。由此得到的量可以直接显示为连续曲线，无需插值。所考虑的总电力差异范围涵盖历史绿色电力缺口以及整个可行的绿色电力盈余区域。

历史绿色电力缺口的内在 p 值。获得式（3.1）中每个目标功率差对应的玻尔兹曼假设分布后，$\mathfrak{h}$ 我们计算式（3.2）中观测到的绿色电力缺口的内在 $\mathfrak{p}$ 值。正如第 3.2 节所述，该值近似为自由度为 1 的累积分布函数 $F_{\chi^2_1}$，其 χ^2 检验统计量为

$$t = 2N \min_{\mathfrak{p} \in \mathcal{A}} \mathrm{D}(\mathfrak{p} \parallel \mathfrak{h}) \,. \tag{4.2}$$

此处，$\mathcal{A}$ 表示支持历史观测绿色电力缺口的分布线性族，与历史存储库族类似。如果 $\mathfrak{h}$为所有观测到的项目概况赋予严格正权重，则检验统计量可简化为 $t = 2N\mathrm{D}(\mathfrak{f} \parallel \mathfrak{h})$。如果$\mathfrak{h}$ 为任何观测到的项目概况赋予零权重，则历史存储库的抽样概率为零，因此在组合假设下它是不可能发生的，而非仅仅不太可能。

图1 显示了由 $1 - F_{\chi^2_1}(t)$ 估计的结果p值剖面，涵盖所考察的规定功率差范围，包括历史绿色电力缺口（其p值显然为1），以及完整可行的绿色电力盈余区域。我们确定了历史基线周围的一个可信区域，其中包含这样一些假设：在规定的$1/N$ 阈值下，历史绿色电力缺口并非极不可能，因此不足以拒绝这些假设。这个区域可被视为置信区域的非公理化对应物，但不应将其解读为表明哪项政策曾在历史上实施或达成一致。它仅确定了在指定结构条件下，与历史基线在统计上相容的绿色电力差异。

请注意，可信区域完全落在绿色电力差异为负值的范围内。相反，在规定的电力差异为正值的整个绿色主导区域中，内在的p值始终远低于接受阈值。[5] 因此，按照假设检验的标准逻辑，我们必须得出结论：在 $1/N$ 阈值下，历史绿色电力缺口与所有涉及绿色电力盈余的受检组合假设在统计上均不相容。换言之，我们可以拒绝这一

[5] 在我们的应用中，反事实概率趋近于零的极限情形出现在更加严格的绿色电力盈余下，此时单个项目配置在达到零化石能源边界之前就已经获得零权重。在零化石能源边界处，所有棕色项目配置都必然被淘汰。无论如何，早在达到此类极限情形之前，p值就已经受到指数级抑制，远低于任何合理的接受阈值。

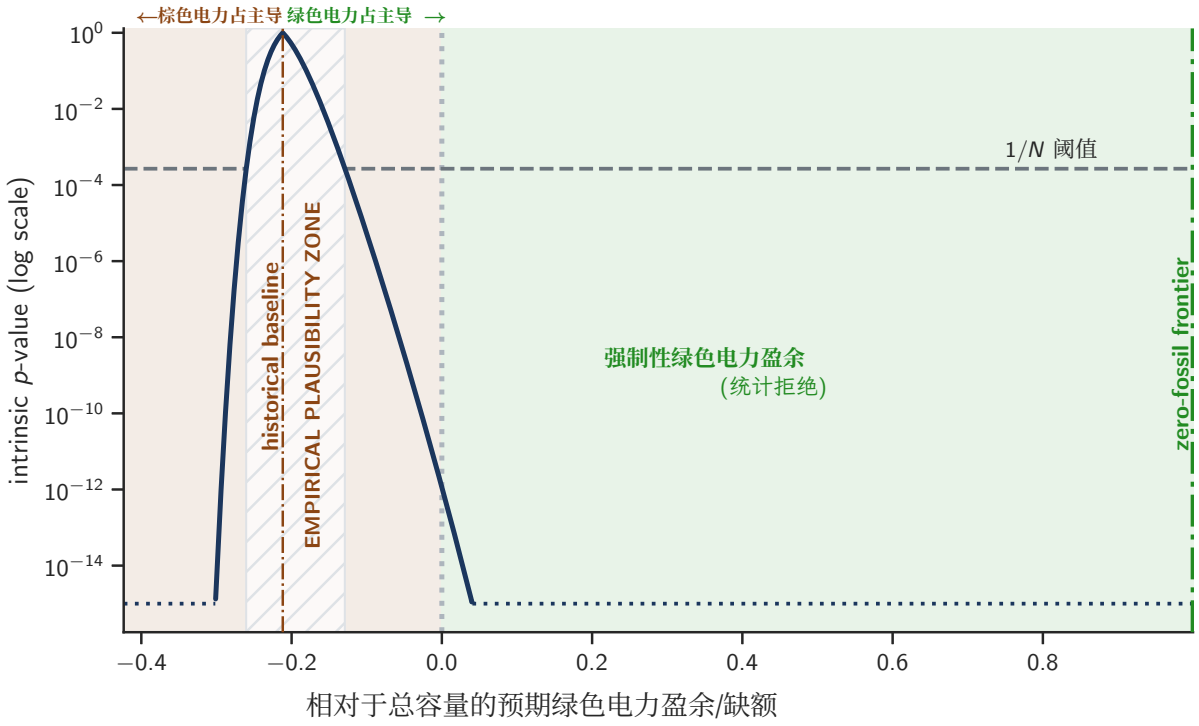


图 1：**内在 p 值在目标电力差异下的分布**。连续曲线描绘了在所考虑的电力差异范围内，对于每种假设反事实，获得历史绿色电力缺口（−22 425.6 兆瓦）或更极端的绿色与棕色电力差异的内在 p 值的半解析近似。计算得出的p 值低于标准机器精度时，以水平虚线表示，以避免在对数尺度上显示呈指数级 N 抑制的波动。

假设：在指定的结构性条件下，强制性绿色电力盈余曾系统性地推动历史项目库的形成。

根据第3.3节，这项统计评估具有直接的政策解读。在所考察的结构条件下，观察到的绿色电力缺口不能仅仅被视为在原本受绿色强制要求的项目形成过程中，由行政惯性及其他背景因素所引发的“偶然”波动。要实现绿色主导，不能依赖历史模式的延续；相反，强制规定的绿色目标要求有意识地改变这种模式。

此时，我们将阐明决定本次统计评估范围及其在当前情境下解读方式的要点。

- **基于抽样的表述：** 内禀的p值定义为仓库族上的累积抽样概率。这自然形成了一种双侧表述：只要某个仓库族的粗粒化抽样概率不高于历史存储库族，该仓库族就会被纳入，而不论其电力差异是处于亏缺侧还是盈余侧。对于当前的一维检验，原则上也可以构建一种针对绿色电力缺口的单侧表述，而不改变假设检验的结论。尽管如此，我们仍保留基于抽样的定义，因为它提供了一种统一的表述

内禀p值，尤其是在同时检验多个强制条件时。

- **严格的拒绝阈值：**我们采用随1/N缩放的拒绝阈值，而非实证研究中常用的传统固定5%阈值。这种固定阈值在很大程度上是一种历史惯例，主要归因于Fisher针对相对较小样本所提出的显著性水平。某一国际组织的历史存储库包含数千个项目。当从假设反事实所对应的如此庞大的替代存储库中抽样时，更为“极端”的偏离可能具有不可忽略的概率。更为严格的阈值（例如1/N）将基于抽样的p值的这一特征纳入考量。重要的是，上文所报告的结论并不依赖于这一严格选择，因为在整个绿色主导区域内计算得到的p值仍远低于任何传统的、高得多的阈值。

- **整合的历史时间跨度：**我们的分析将1984年至2024年这四十年期间所有包含有效、可归类为绿色的项目的历史记录整合为单一历史存储库。这一延长的时间跨度降低了个别年份内可能出现的波动影响，并增加了统计评估可用的样本量。它还提供了更多观测到的项目概况。特别是，历史存储库中所包含的更广范围的绿色与棕色电力差异，使观测到的项目概况中可行的组合假设范围，比在大幅缩短的历史时期内所能实现的范围更广。总体而言，我们的统计评估关注的是整个四十年时间跨度内历史存储库的总体形成过程，而非任何特定年份的波动。

- **结构条件：**假设检验是在总铭牌容量和水电容量固定的结构条件下进行的，这两项容量均根据历史水平进行校准，详见第 3.1节。这些条件对反事实情景施加了物理和运行约束。尤其是，改变总铭牌容量会改变历史承诺的规模，而水电容量则受到技术、运行、物理和环境方面的约束，适宜水资源的可用性也受到这些因素限制。尽管我们对绿色电力盈余假设的统计拒绝受到这些结构条件的约束，但在政策问题或现有信息需要时，玻尔兹曼框架允许解除、替换或扩展这些条件。

投资分配的转变。在确立历史绿色电力缺口在统计上与任何强制规定的绿色电力盈余不相容之后，我们现在转而讨论绿色政策的财政影响。继统计评估得出这一重要结果后，相关问题已不再是绿色转型能否自然地从行政惯性以及塑造历史存储库的更广泛现实世界进程中出现。由于这种自然出现的可能性极低，因此必须采取主动干预来改变

这一历史模式。事前政策探索现在涉及在指定的结构条件和治理原则下，伴随绿色转型而来的投资重新分配。

为考察财政后果，我们研究第2节所述的、描述项目总投资的X可观测量。1请注意，这一可观测量并未纳入我们在第3节1中构建的绿色强制要求组合假设。然而，对其进行分析与绿色预算的原则和框架直接相关；在绿色预算中，预算分配被用作推进环境目标的政策工具。在针对每个规定的电力差异构建的假设反事实中，绿色转型导致的投资重新配置的总体效应通过X的矩加以捕捉。

对于每个目标绿色与棕色电力差异，我们追踪玻尔兹曼假设反事实中分配给各技术标签的预期总投资份额。对于给定的技术类别s，该份额表示为

$$\sum_x x\, h(x,s), \qquad s \in \{\texttt{green},\texttt{brown}\}, \tag{4.3}$$

相对于总体预期项目成本，

$$\sum_x x\, h(x)\ . \tag{4.4}$$

这些表达式中的权重是通过公式（4.1）中的信息散度最小化得到的玻尔兹曼假设分布的权重。根据构造，对于所考虑的每个目标电力差异，绿色与棕色投资份额之和均为一（100%）。

图2 中的轨迹揭示了标记投资分配与标记铭牌容量之间的不对称性。在历史基线下，大部分投资都投向棕色项目。在投资交叉点，即绿色项目与棕色项目的支出相等时，棕色项目所占的总容量份额仍然更大（电力差异为负）。因此，沿着迈向绿色主导的转型路径，投资平价先于电力平价出现。在玻尔兹曼假设反事实中，要形成绿色电力盈余，分配给绿色项目的预期投资必须略高于一半。

公式（4。4）中的每个项目预期成本，在不同规定的电力差异下呈现非单调轨迹。为便于与历史基线比较，我们将这一数值按其历史值进行标准化，该历史值由公式（ ɦ = ƒ（4。4）得出。从历史基线进入绿色区域时，预期项目成本最初会上升至比历史基线高约 3%。然而，在绿色电力主导区域更深处，这一趋势会逆转，预期成本大幅下降；在零化石能源边界附近，降至比历史基线低约 9%。因此，在反事实构造中，日益严格的绿色电力授权并不意味着预期项目成本会单调增加。

综上所述，这些结果为绿色预算编制提供了两点直接相关的考量。首先，投资分配平价并不意味着标签电力平价。在其他条件不变的情况下，

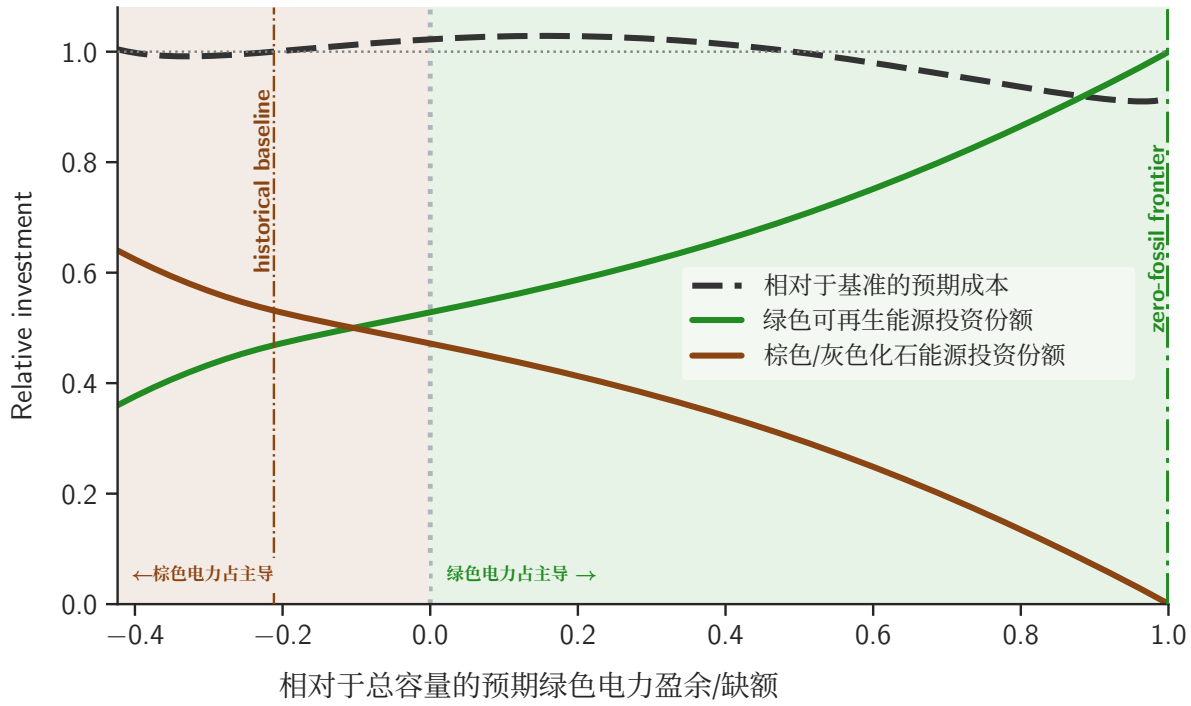


图2：**目标电力差异下的投资分配与项目成本**。连续曲线描绘了分配给绿色与棕色技术标签的预期投资份额。投资份额在电力平价（零电力差异）之前的棕色主导区域相交。虚线曲线描绘了预期项目成本（按其历史基线归一化），该成本随目标电力差异呈非单调变化。

从历史基线开始进行最低限度的转型时，绿色项目与棕色项目之间的投资分配会先达到均等，之后才会实现二者电力贡献的平衡。其次，纯粹从财政角度来看，较低的绿色电力盈余似乎并不可取，而追求更高的绿色电力盈余似乎更有利，因为这能让绿色转型将预期项目成本降至历史基线以下。换言之，仅从预期项目成本的变化轨迹来看，政策制定者可以更积极地推动绿色转型。

图1 和2 自然共享相同的格式模式。统计评估与财政分析依赖于同一反事实构造，前者评估历史基线在规定的电力差异下的合理性，后者量化与最小规定转型相关的投资分配变化。在技术层面，同一玻尔兹曼假设分布 $\mathfrak{h}$，针对式（3.1）中的每一电力差异进行求值，并同时进入式（4.2）中的检验统计量以及式（4.3）和（4.4）中的预算指标。

4.2 绿色转型的进一步影响

现在，我们转而讨论假设反事实的几个方面，这些方面涉及项目构成、技术分配和发展相关差距。更多分析

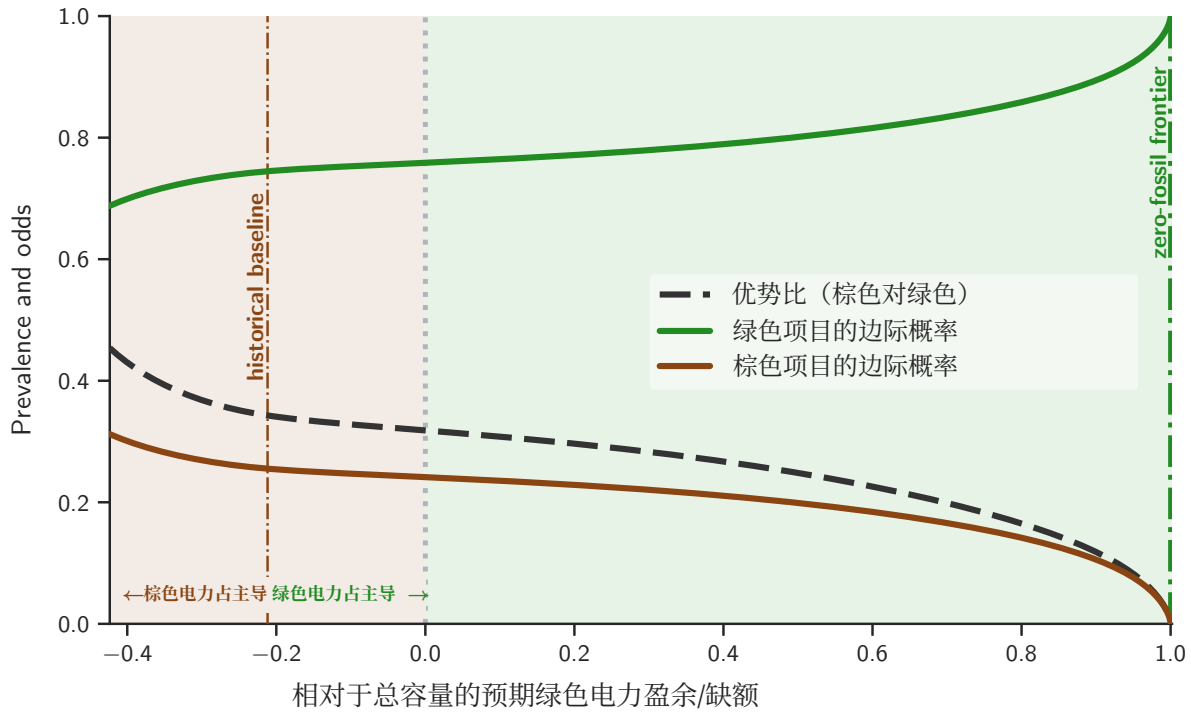


图3：**不同目标电力差异下的项目占比**。这些曲线描绘了绿色与棕色技术标签在各目标电力差异下的预期相对频率（反事实概率），以及相应的选择优势比。

见附录A。3。

绿色转型中的项目占比。我们首先考察绿色标签项目与棕色标签项目在各目标电力差异下的占比。在图3中，$h(s)$ 针对$s \in \{\texttt{green}, \texttt{brown/grey fossil}\}$ 的轨迹展示了假设反事实的标签构成如何随规定的电力差异变化而改变。值得注意的是，从历史绿色电力缺口过渡到精确的绿棕功率平衡，仅需将绿色项目的相对频率提高1.4%。因此，向电力平衡的转型最初只需对带标签技术类别的占比作出适度调整即可实现。

占比轨迹的相对稳定性反映了最小偏离反事实的正则化效应。I投影并非优先将单个项目概况的概率设为零以实现规定的电力差异，而是在观测到的概况之间重新分配概率质量，并在施加条件允许的范围内尽可能保留历史构成。只有当该规定本身要求棕色电力为零、即处于零化石能源边界时，棕色项目才会被完全消除。

细粒度资源配置与资本效率。绿色标记是绿色预算的重要前提，但它并未揭示支出如何在各项绿色技术之间分配，也未说明这种分配伴随哪些经济特征。事前评估不应止于统计被归类为绿色的支出总额，还应考察这些方面。

我们的反事实构造提供了这一额外的分析层次，无需直接设定任何技术特定的政策偏好。由于玻尔兹曼假设分布为完整项目概况赋予反事实概率，因此，历史项目库中任何可计算的量，同样可以在相应的假设反事实中进行评估。具体而言，我们在按公式（˜2.5）引入的细粒度变量 S分层的下游分析中，利用Y和X可观测量，考察财政和经济后果。

这些后果通过细粒度总投资（X的一阶矩）进行量化，

$$\sum_{x} x\, p(x, s = \texttt{green}, \tilde{s}) \,, \tag{4.5}$$

我们将其按绿色总投资进行归一化

$$\sum_{x} x\, p(x, s = \texttt{green}) = \sum_{\tilde{s} \neq \texttt{N/A}} \sum_{x} x\, p(x, s = \texttt{green}, \tilde{s}) \tag{4.6}$$

以获得图4上图所示的归一化绿色投资份额。这确保上图中的五条曲线在每个目标功率差下均加总为 1。作为经济度量型指标，我们采用图4下图所示的资本效率系数，其定义为某一绿色技术类别的单位总投资所对应的铭牌容量。在玻尔兹曼假设反事实中，该系数等于细粒度总容量（Y的一阶矩）与

$$\sum_{y} y\, p(y, s = \texttt{green}, \tilde{s}) \,, \tag{4.7}$$

公式（4.5）中的细粒度总投资之比。

在图4中，轨迹揭示了绿色投资分配如何随规定的功率差而变化。在历史基线下，水电项目占绿色总投资的份额最大；而随着向绿色电力主导转型，投资分配逐步转向陆上风电。其余绿色技术在整个转型过程中所占份额较小，仅呈边缘波动。具体而言，太阳能机群在整个转型过程中均获得了相当可观的绿色投资份额，而地热和生物能源与废弃物的资金份额始终很低。

资本效率轨迹为财政重新分配提供了互补的经济视角。陆上风电的绿色投资份额不断上升，同时其资本效率系数也显著提高；相比之下，水电项目的相应系数则相对稳定。在强制绿色转型过程中，其他技术的资本效率系数较低，变化也不那么明显。

式（3.4）中的固定水电容量条件对于解读这些轨迹至关重要。如果不加约束，在严格的绿色电力盈余条件下，水电将在玻尔兹曼假设反事实中占据主导地位，无论是绿色投资还是资本效率系数方面都是如此。这种主导地位并非方法论假象，而是

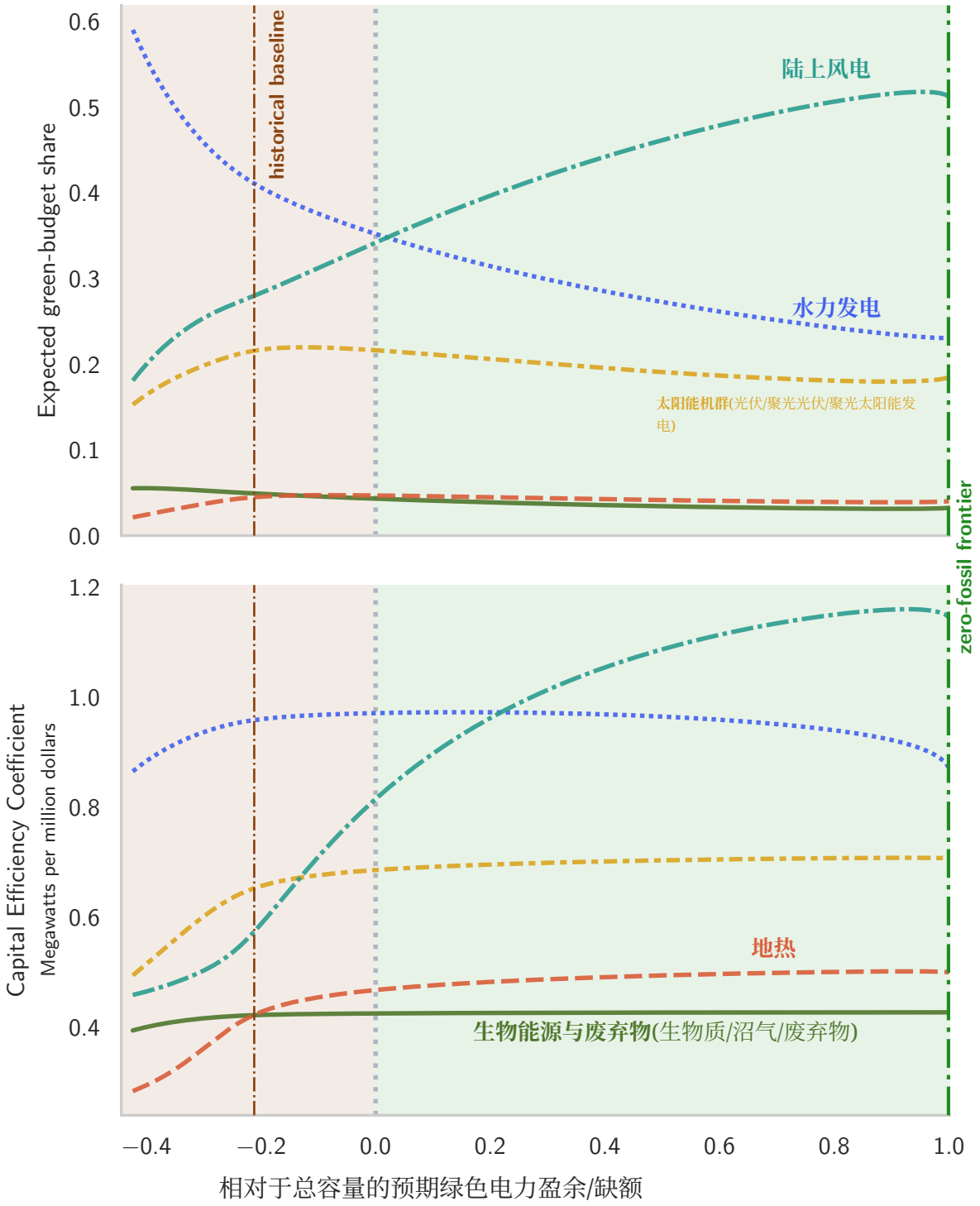


图4：**不同目标功率差下的绿色投资与资本效率**。上图展示了五种绿色标记技术的绿色投资份额，显示投资主导地位如何从水电转向陆上风电。下图展示了相应的资本效率系数，显示陆上风电的资本效率大幅提升，而其余技术的轨迹则相对稳定。两个面板采用相同的配色方案。

不过是水力发电规模的结果，水坝构成了大型基础设施项目，提供了异常巨大的铭牌容量。如下文式（3.4）所解释的，我们事前分析的相关外生约束是，这一历史水电规模无法无限外推至未来。因此，我们选择在整个绿色转型过程中保持总水电容量不变。[6]

总而言之，细粒度的财政与经济轨迹源于对组合假设条件的最小调适，这些条件既不涉及财政可观测量X，也不˜涉及细粒度可观测量S。由此，玻尔兹曼假设反事实将陆上风电，其次是水电，然后是太阳能机组，确定为在指定外生约束下日益严格的绿色转型的驱动技术。

国际差异。为考察投资效率与发展相关的维度，我们分析按次级可观测量分层的资本效率系数$Z_{3,}$，其对应于世界银行国际开发协会（IDA）资格状况。对于每个发展类别$z_{3,}$，该系数（参见式（4.7）和式（4.5））由总铭牌容量与总投资之比给出：

$$\sum_{y} y\, h(y|z_3) \Big/ \sum_{x} x\, h(x|z_3)\ . \tag{4.8}$$

回顾一下，$\mathfrak{h}$已针对每个目标电力差异完成评估。

图5 显示，三个发展类别之间的资本效率差异持续存在。非国际开发协会国家的资本效率系数最高，随着目标电力差异的变化而波动，并在绿色主导区域内达到局部最低点，但仍高于每百万美元（不变价）的0.88兆瓦每百万不变价美元。相比之下，低收入国际开发协会资格国家和混合融资国家的轨迹要么几乎停滞不前，要么持续下降；随着反事实情景逼近零化石能源边界，这两类国家的数值均降至每百万美元（不变价）0.6兆瓦每百万不变价美元以下。

分层结果对于解读观察到的资本效率差距具有重要意义。国际开发协会地位根据各国的融资和发展条件对其进行区分：世界银行依据人均收入较低、按市场条件借款的信用资质有限等标准，将符合国际开发协会资格的国家归类；混合融资国家既符合国际开发协会资格，又能在一定程度上获得国际复兴开发银行融资；而在这一分类中，非国际开发协会国家仅有资格获得国际复兴开发银行贷款 [26]。假设这些差异可作为底层基础设施和投资基础等方面差异的代理指标，那么观察到的三类国家资本效率历史差异表明，各国将投入资本转化为铭牌容量的经济能力存在显著的分配差异。

由于我们没有直接控制这些差异，玻尔兹曼假设下的反事实情景会将历史资本效率差距贯穿整个强制转型过程传播开来。更重要的是，三组之间的总体差距变得愈发明

[6] 从纯粹的方法论角度来看，玻尔兹曼框架允许将水力发电总量设定为任意值——只要该值有观测到的铭牌容量作为依据，即可偏离历史基线。

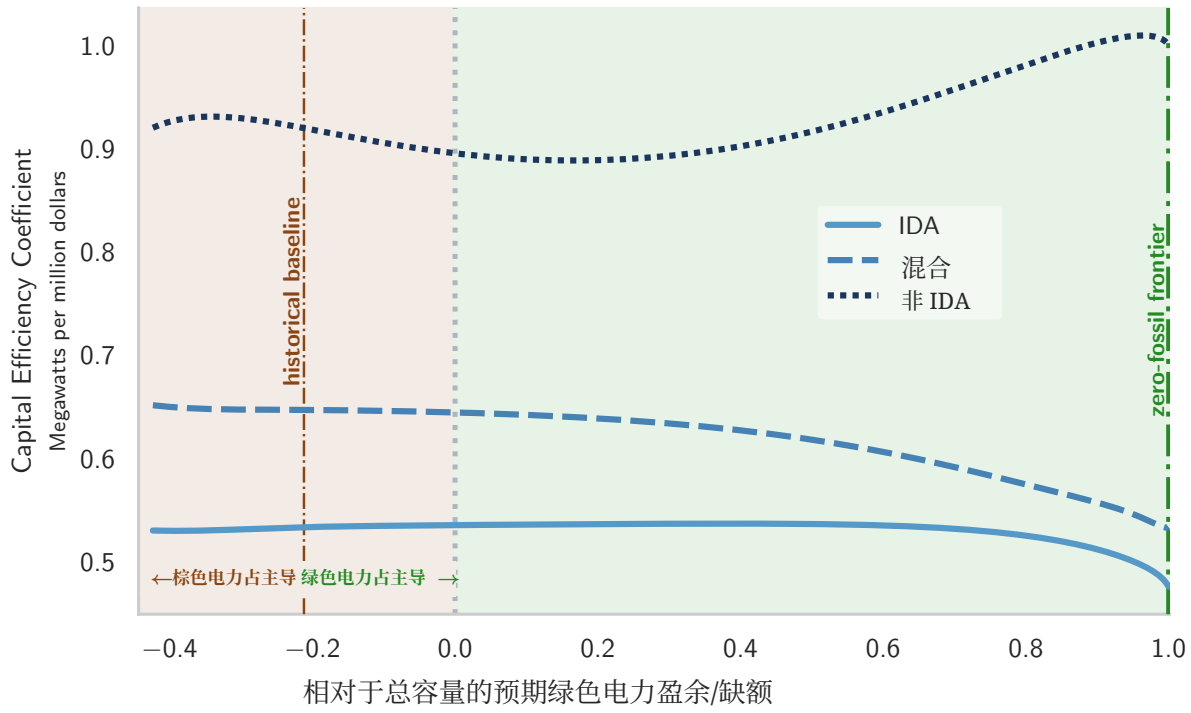


图 5： **按 IDA 地位划分的资本效率**。图中展示的是总铭牌容量与总投资之比，其定义见式（4。8），并按 Z_3 在各目标电力差异下进行分层。

显，因为项目库正逼近零化石能源边界。从政策视角来看，这种趋势尤其令人担忧。如果在没有附加结构条件的情况下施加日益严格的绿色电力盈余要求，转型可能会加剧而非缩小投入资本转化为铭牌容量的效率差异。

下游分析的最后部分强调，仅有绿色电力指令并不能解决各国资本效率方面观察到的差异。为避免绿色转型加剧现有差距，必须在政策目标之外补充旨在解决资本效率历史差异的结构性条件。幸运的是，玻尔兹曼框架可以无缝扩展组合假设，将此类以公平为导向的条件纳入其中。

5 结论与展望

5.1 结论

我们采取一种有别于传统统计学的正交方法，通过明确构建玻尔兹曼假设反事实，为政策目标的事前影响评估提供了一个易于处理的解决方案。该反事实情景完全由历史项目库、政策目标以及政策制定者认为直接相关的任何约束条件所决定。

具体而言，我们考虑这样一类政策目标：其组合假设可以通过约束相应广义矩的聚合条件来表达。我们关注这样一些情形：对于支配项目选择和资料库形成的实际过程，人们一无所知，也不作任何其他假设。在先前的研究中，我们通过大N 鞍点技术证明，这种最大限度不可知的抽样过程的物理规律，提供了柯尔莫哥洛夫框架试图通过公理化推导得出的要素。特别是，对抽样概率进行粗粒化，不可避免地会在最传统的费希尔意义上产生一个近似易于处理的内在p值。

秉持统计物理学视角，我们表明，一个包含任意模态变量的历史存储库，加上政策目标的组合假设，就足以唯一构造玻尔兹曼假设反事实。仅将分析限定于观测到的项目概况时，这种唯一性源于内在的p值的认知一致性。玻尔兹曼假设反事实通过内在的p值对政策目标进行统计评估，同时也能评估原始历史存储库中可计算的任何下游财政或社会经济指标。值得注意的是，由认知一致性所选择的这一假设反事实的最小偏离性质，为评估比例原则提供了坚实基础。

每当大N 状态适用时，任何声称仅依据历史基准和组合假设得出的、关于政策目标且自洽的统计或下游事前评估，无论采用何种统计框架，都必须与我们的主阶结果一致。根据玻尔兹曼假设反事实的构造，在这一状态下，任何主阶层面的不一致都必然表明，该备择假设评估有意或无意地通过纳入有关项目选择、其他行政方面等的假设，“修正”了历史存储库的构成。只要明确承认这些额外假设，它们本身并无不可取之处。

作为原理性证明，我们使用世界银行公开数据，在绿色强制要求所规定的绿色与棕色电力差异的可行范围内，构造了玻尔兹曼假设反事实。对于任何可行的绿色电力盈余，其内在 p 值都极低，这意味着，如果当时已实施绿色强制要求，单凭非系统性现实因素就造成历史绿色电力缺口，极不可能发生[7] ，前提是这些因素会影响资料库的形成。因此，根据统计假设检验的传统说法，我们得出结论：当时并未实施绿色强制要求，因此需要主动干预，以改变历史模式，从而实现任何被认为理想的绿色电力盈余水平。

下游分析显示，绿色与棕色投资达到平价早于绿色与棕色电力达到平价，而总平均项目成本则随着规定的功率差呈现非单调轨迹。私人所有权和项目年份相对稳定。在技术层面，玻尔兹曼假设反事实指出，陆上风电最终将成为推动绿色转型的力量。区域投资分配总体保持稳定，规定的调整主要通过地区内部将资金从棕色电力重新分配至绿色电力来吸收。资本效率差异

[7] 在大尺度描述中，非系统性现实因素可由玻尔兹曼阐释性思想实验中的抽样不确定性有效捕捉。这种不确定性是柯尔莫哥洛夫所提出的随机性的一种物理对应。

在 IDA、混合型和非 IDA 经济体之间持续存在，并随着绿色强制要求日益严格而扩大。

世界银行的应用展示了我们的框架带来的三项实际影响。首先，内在p值用于评估历史基线与强制实施绿色转型之间的统计兼容性，而下游的财政和社会经济后果则为后续政策决策提供依据。其次，同一玻尔兹曼假设反事实可评估每种可行的联合假设所对应的这些下游后果，而无需建立新的统计或计量经济模型。第三，玻尔兹曼假设分布的高效计算以及内在p值的闭式近似，使得数千项规定目标能够在数秒内完成评估。

由此，所提出的框架为事前绿色预算编制提供了全面的量化支持。这与欧盟《绿色预算参考框架》对政策评估、透明度和问责制的强调相一致 [27]。实现强制目标在财政、经济、技术和地理维度上的影响，可为评估政策与适用的财政和经济约束条件是否相容提供依据。此外，此类量化评估有助于制定绿色中期财政框架；此类框架旨在将气候政策的经济和财政后果纳入战略性公共财政规划 [1]。

受统计物理学启发，我们的大规模推理带来的优势在于：在同一构建框架下，我们能够迅速评估绿色强制要求组合假设下历史赤字的下游后果及其统计上的（不）可能性。这种自洽的评估从不依赖于对项目库形成过程的总体模型或微观物理模型，因此可以直接归因于历史存储库和政策制定者所提出的组合假设。

5.2 展望

尽管本文以从世界银行项目库提取的数据集为例进行了演示，玻尔兹曼构建方法仍可广泛应用。它同样适用于国家部委或其他公共当局维护的项目库，从而能够在单个国家或国家群体（如欧盟成员国）层面开展评估。反事实权重可通过标准的优化牛顿-拉夫森算法低成本计算，而内在的p值则可使用相应的 χ^2 统计量近似得出，无需进行拒绝抽样。因此，随着项目的增加、技术标签的重新分类或绿色强制要求的更新，玻尔兹曼假设反事实都可以高效地重新计算。

当历史存储库包含时间观测量时，连续的历史时间窗口便可提供回溯性的分析方向。给定一项指令后，可根据较早时间窗口构建玻尔兹曼假设反事实，并将其与较晚时间窗口中观测到的、在该指令下实现的下游结果进行比较。由此产生的差异有助于识别还需要哪些额外的政策目标或结构条件，才能使较早的反事实重现较晚的结果；这也为追溯观测期内实际的政策干预或其他系统性影响提供了一种方法。

这　构建从根本上与领域无关；特别是，ori的模态　g-

inal属性在一般抽样的有效描述中变得无关紧要。这使得它适用于任何须在最小重组下将低维社会目标与高维资产组合、项目或其他单元相协调之处。潜在应用包括国际人道主义响应中的资源分配，包括移民和难民危机，以及选拔程序中的人口代表性，例如国家和区域劳动力结构中持续存在的性别失衡 [28]。线性代数表述也能容纳多重政策目标，例如将公平性条件与绿色目标一并纳入。

一项方法论扩展涉及政策目标的形式。作为这种正交方法的首个见解，我们处理了组合假设：这些假设将可观测量函数的期望值约束为指定值，从而对广义矩施加线性条件。将这一构造扩展到期望值的非线性条件，可以容纳广义矩之间更复杂的代数关系，进而处理更复杂的政策问题。

另一项扩展涉及资料库形成。本研究采用刻意保持不可知的思路，将这一过程实质上视为多项分布，未明的现实世界影响则纳入通用抽样不确定性之中。如果实际项目选择中的系统性特征已知且有独立依据，则可通过广义矩条件或多项抽样以外的广义抽样方案 [29]将其纳入。这使得有关资料库形成的额外信息能够以受控方式进入反事实构造，而无须对其背后的社会经济、技术或行政机制建立完整模型。

数据可用性

本研究使用的数据可从世界银行基础设施私人参与（基础设施私人参与）数据库公开获取。随附的 Jupyter 笔记本https://gitlab.uni-marburg.de/fb20/ag-chung/greenboltzmann提供了完整的计算工作流程，涵盖数据处理、历史存储库的确定、玻尔兹曼假设反事实的构造以及后续分析。其中包括计算论文图表所用的内在 p 值和下游指标。

致谢

K.S. 和 G.F.L. 的研究工作得到让·莫内讲席“欧盟经济与财政治理框架及透明度”（EUFiscalGov）的支持。该讲席由欧盟通过伊拉斯谟+计划资助（项目编号：101239868）。

Funded by the European Union

免责声明

所表达的观点和意见仅代表作者本人，并不一定反映欧盟或欧洲教育与文化执行局的观点和意见。欧盟和资助机构均不对这些观点和意见承担责任。

参考文献

[1] F. Caselli、A. Lagerborg 和 P. Medas。绿色财政规则？挑战与政策备选方案。国际货币基金组织工作论文 2024/125。华盛顿特区：国际货币基金组织，2024年6月。[2] 欧盟委员会。《更好的监管》。访问日期：2026年10月6日。2026年。网址：https://commission.europa.eu/law/law-making-process/better-regulation_en。[3] 经济合作与发展组织。《关于监管政策与治理的理事会建议》。巴黎：经济合作与发展组织出版社，2012年。ISBN：9789264209022。[4] J. Hertin 等。《让政策混乱趋于理性？事前政策评估与政策过程中知识的运用》。《环境与规划A》41.5（2009），第1185–1200页。[5] T. Grüne-Yanoff 和 C. Marchionni。《模型多元主义中的模型选择建模》。《经济学方法论杂志》25.3（2018），第265–275页。[6] W. A. Brock、S. N. Durlauf 和 K. D. West。《不确定的经济环境中的政策评估》。《布鲁金斯经济活动论文集》2003.1（2003），第235–322页。[7] S. N. Durlauf。《复杂性、经济学与公共政策》。《政治学、哲学与经济学》11.1（2012），第45–75页。[8] 欧盟委员会。《更好的监管工具箱》。2025年12月1日。网址：https://commission.europa.eu/law/law-making-process/better-regulation/better-regulation-guidelines-and-toolbox/better-regulation-toolbox_en。[9] S. N. Durlauf。《社会经济行为的统计力学方法》。美国国家经济研究局技术工作论文 0203。美国国家经济研究局，1996年9月。[10] S. N. Durlauf。《统计力学如何为社会科学作出贡献？》。《美国国家科学院院刊》96.19（1999年9月），第10582–10584页。[11] B. A. Furtado、M. A. Fuentes 和 C. J. Tessone。《政策建模及其应用：最新进展与展望》。《复杂性》2019（2019），第5041681页。[12] C. Castellano、S. Fortunato 和 V. Loreto。《社会动力学的统计物理学》。《现代物理评论》81.2（2009），第591–646页。[13] A. De Martino 和 M. Marsili。《异质主体社会经济系统的统计力学》。《物理学杂志A：数学与通论》39.43（2006），R465–R540。[14] L. Boltzmann。“关于若干机械热理论问题的进一步评论”。德文。载于：《科学论文集》。F. Hasenöhrl 编。剑桥大学出版社，1878年8月，第250–288页。ISBN：978-1-108-05280-1。网址：https://www.cambridge.org/core/product/identifier/CBO9781139381437A018/type/book_part。[15] O. Loukas。《抽样的玻尔兹曼结构：内在 p 值及其涌现的闭式表达》。2026年。arXiv：2608。14608 [math.ST].[16] A. Gelman 和 E. Loken。《〈分岔路径花园〉：为什么多重比较可能会造成问题，即使不存在“钓鱼式研究”或“P值操纵”，且研究假设是事先提出的》。哥伦比亚大学统计系（2013）。工作论文。[17] 世界银行。基础设施私人参与（PPI）。https://ppi.worldbank.org/en/customquery。访问日期：2026年9月15日。[18] 世界银行。基础设施私人参与（PPI）：术语表。https://ppi.worldbank.org/en/methodology/glossary。访问日期：2026年9月15日。[19] O. Loukas 和 H.-R. Chung。《科学的完全经验基础》。2024年。arXiv：2410.19866[physics.data-an]。

[20] I. Csiszar。《概率分布的 I-散度几何与最小化问题》。《概率年刊》3.1（1975年2月）。出版方：数理统计学会，第146–158页。网址：https://projecteuclid.org/journals/annals- of-probability/volume- 3/issue- 1/I- Divergence-Geometry-of-Probability-Distributions-and-Minimization-Problems/10.1214/aop/1176996454.full.[21] B. Efron。《自助法：再看刀切法》。《统计年刊》7.1（1979年1月）。出版方：数理统计学会，第1–26页。网址：https://projecteuclid.org/journals/annals-of-statistics/volume-7/issue-1/Bootstrap-Methods-Another-Look-at-the-Jackknife/10.1214/aos/1176344552.full（访问日期：2024年12月12日）。[22] 经济合作与发展组织。农业促进土壤碳封存的政策选项。技术报告。COM/ TAD/CA/ENV/EPOC(2020)8/FINAL。经济合作与发展组织，2022年1月。网址：https://one.oecd.org/document/COM/TAD/CA/ENV/EPOC(2020)8/FINAL/en/pdf。[23] P. Larrue、P. Tõnurist 和 D. Jonason。《使命导向型创新政策的监测与评估：从理论到实践》。技术报告。2024/09。巴黎：经合组织出版部，2024年。网址：https://doi.org/10.1787/5e4c3204-en。[24] 欧盟。《欧洲联盟条约》合并版，第5条。《欧盟官方公报》C 326，第18页。2012年。网址：https://eur-lex.europa.eu/eli/treaty/teu_2012/ art_5/oj（访问日期：2026年10月12日）。[25] R. L. Wasserstein 和 N. A. Lazar。美国统计学会关于p值的声明：背景、过程与目的。《美国统计学家》70.2（2016年），第129–133页。[26] 世界银行集团。基础设施私人参与（PPI）的国家分类。2020年。网址：https://ppi.worldbank.org/en/methodology/country-classification（访问日期：2026年9月28日）。[27] 欧盟委员会。《欧盟绿色预算编制参考框架》。技术报告。欧盟委员会，2022年。[28] O. Loukas 和 H.-R. Chung。《人口统计均等：减轻真实世界数据中的偏差》。2023年。arXiv：2309.17347 [cs.LG].[29] R. K. Niven 和 M. Grendar。广义经典统计、量子统计与中间统计及波利亚瓮模型。《物理快报A》373.6（2009年），第621–626页。[30] I. Csiszár 和 P. C. Shields。《信息论与统计学：教程》。《通信与信息论基础与趋势》1.4（2004年12月）。出版方：Now Publishers, Inc.，第417–528页。ISSN：1567-2190，1567-2328。网址：https://www.nowpublishers.com/article/Details/CIT- 004。[31] Y. Kitamura 和 M. Stutzer。广义矩估计法的信息论备择方法。《计量经济学》65.4（1997年），第861–874页。[32] G. W. Imbens、R. H. Spady 和 P. Johnson。矩条件模型推断的信息论方法。《计量经济学》66.2（1998年），第333–357页。[33] S. M. Schennach。基于指数倾斜经验似然的点估计。《统计年刊》35.2（2007年），第634–672页。[34] E. T. Jaynes。“熵最大值处分布的集中（1979）”。载于：E. T. Jaynes：《E. T. Jaynes：概率、统计与统计物理论文集》。R. D. Rosenkrantz 编。Synthese文库。多德雷赫特：Springer Netherlands，1989年，第315–336页。

A 附录

A.1 玻尔兹曼统计框架概要

表1：**公共治理中的统计物理学跨学科术语**。在有效独立抽样普适类中，对备选项目库的选择概率进行粗粒化后，统计结构便会显现。示意图见图7。

概念／构建	描述／说明
样本量 N	历史记录的抽样单位数量，用于定义 有限样本不确定性的尺度，从而其适用性 玻尔兹曼统计物理学视角的。
基础设施项目	它们构成抽样单位，亦即抽象球，位于玻尔兹- 曼瓮之中。 这些是被反复抽取的对象 思想实验以及其观测到的特征所 确定的项目概况。
可观测量	与事前分析相关的选定项目属性，例如 投资预算、私人投资占比、铭牌容量、 技术或地理特征。 可观测量并非 先验地声明为度量型或类别型；其模态取决于 施加于其上的条件形式。
项目概况	项目概况是对单个项目所选可观测量的联合规定。 项目概况编码了可能的组合 用于区分球类型的项目特征组合 在玻尔兹曼 urn 中。
观测到的项目概况	仅纳入可用数据中记录的有效项目概况， 以防止反事实构造外推 至特征不合理的项目。观测到的项 已提供的剖面数量不超过 N。
历史存储库	由 N 个历史观测项目构成的有限集合，以各自不同的项目概况及其对应数量表示。 以其各自不同的项目概况及相应数量表示。
经验 分布 $\hat{f}$	直接从历史存储库中获取，它是分布 其概率由以下相对频率给出： 观测到的项目概况。
反事实 分布 p	一种根据某些因素赋予非负权重的分布 相对于观测到的项目概况的反事实情景。

下一页继续

汇总表——续自上一页

概念／构造	描述／解释
反事实库	一个可由观测到的 项目概况构造出的无限大项目库。其构成由反事实 分布 $\mathbb{p}$。
一般抽样	大样本中项目选择的有效描述 情形下项目选择的有效描述。它体现了项目概况普遍程度中的非系统性波动， 这些波动由抽样不确定性所致，并受到许多未知因素的影响。 这一概念源自玻尔兹曼的思想实验—— 验，其中一组N 个项目是独立抽取的 自反事实项目库。
结构条件	对广义矩规定取值的聚合条件 ，即可观测量函数的期望值，独立于 特定绿色目标。在本应用中， 其中包括固定的总容量与水电容量，以及 与固定库规模相关的归一化。
绿色授权条件	规定强制性绿色电力 盈余的均值差条件。 根据政策目标，可能存在多个强制性 可以同时施加的条件。
组合假设	结构与绿色授权条件的集合。它通过 总量的具体值表达所有直接可控的 政策问题中的维度。
假设分布	满足可行的组合假设的反事实分布 假设
假设反事实	其构成由某一 假设分布描述的反事实库。
替代存储库	规模同为N 的项目库，与历史项 目库相当，从假设反事实中抽样，处于结 构条件之下。它们代表替代历史版本，于 该假设反事实的政策全集之中。
存储库族	替代存储库按其总量进行分组 绿电与褐电功率差。不同的备择假设库 因此可以属于同一家族，只要它们表现出相同的 总功率差以及相同的结构性一般化 矩。

下页续

汇总表——承接上页

概念／构造	描述／说明
历史 存储库族	由历史观察到的绿色所定义的仓库族- 功率亏缺。历史存储库本身是这一 线性族中的一个成员；根据玻尔兹曼假设分 布的构造，它是其中最可能的成员。
玻尔兹曼假设 分布 $\mathfrak{h}$	将 $\mathfrak{f}$ 投影到假设线性族上的 I 投影。可归因性 反事实情景与经验结构之间的关系并未被 组合假设直接规定，但这明确确定了 该投影对应的假设分布。
玻尔兹曼 假设反事实	其构成由……决定的假设反事实 玻尔兹曼假设分布，实现政策制定者的 与历史存储库偏离最小的反事实情景 。 在此反事实情景下，历史存储库是 历史存储库族中最有可能出现的。
内在p值	历史存储库族及 以及所有在通用情形下概率并不更高的项目库族 抽样下，来自玻尔兹曼假设反事实的概率不高于 该族的所有存储库的累积概率。它量化了历史上观测到的绿色电力缺口，或 更极端的电力差异，在特定管理的 绿色电力盈余下出现的可能性。

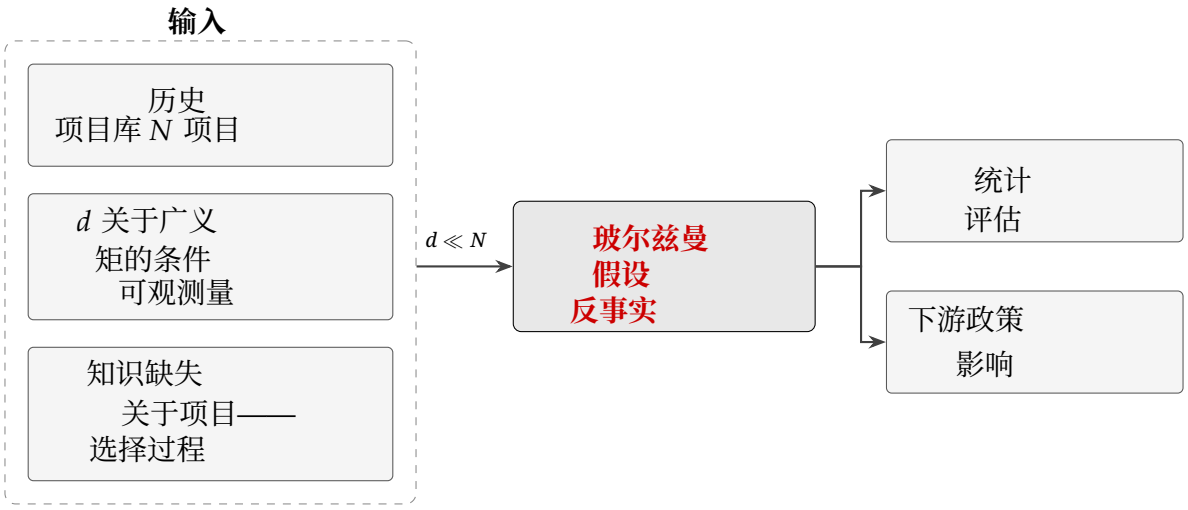


图 6：**玻尔兹曼框架流程图**。从历史存储库中可通过可观测量的广义矩表达的最简政策问题输入，到全面的事前影响评估。

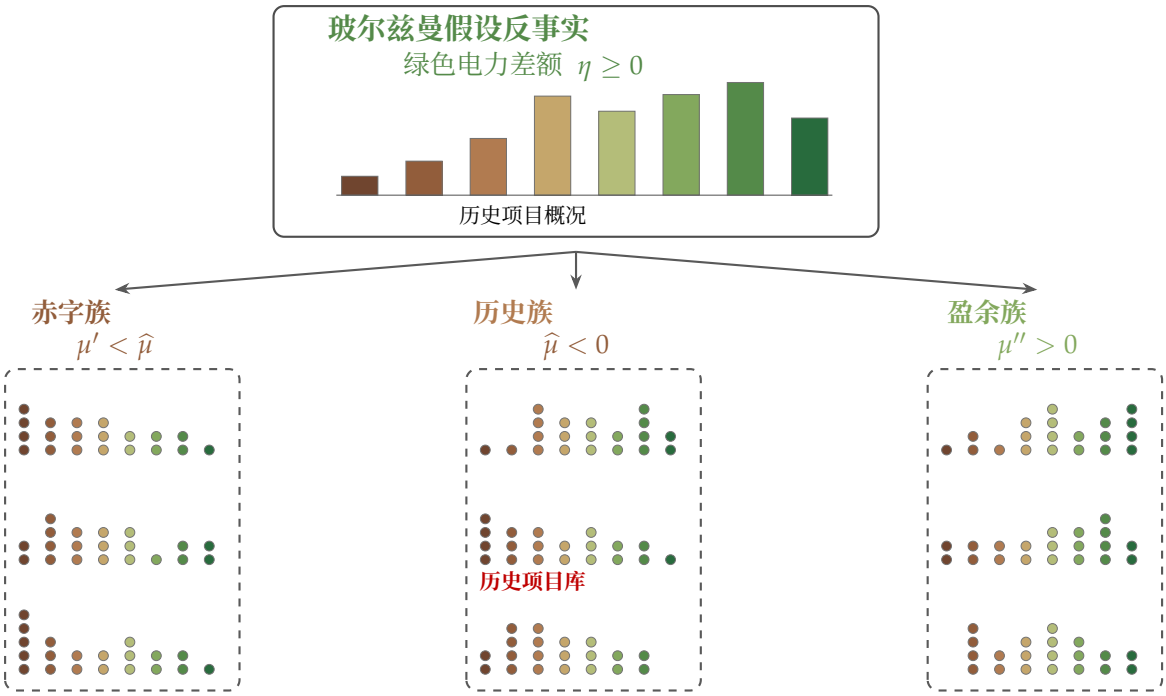


图7：**项目选择有效描述的示意图**。给定样本量N以及进一步的结构条件，我们原则上可以生成历史项目概况上所有组合允许的替代存储库。根据其样本均值 μ，备选项目库被归入各个族。在玻尔兹曼假设反事实下对项目进行有效独立（多项分布）选择时——该反事实包含按满足绿色强制要求之组合假设 η的比例构成的项目概况——替代存储库被赋予概率。将这些概率在存储库族上粗粒化，可得到关于N的领头阶闭式表达式。玻尔兹曼假设反事实的（直方图）构成，是按选择概率对替代存储库加权平均的结果，即使没有任何单一替代项目库满足所强制要求的 $\eta \geq 0$。

A.2 数学背景

由于铭牌容量上的总量不过是可观测量Y的广义矩，式(3.3)与式(3.4)定义了观测到的项目概况上的库（数据集）的结构线性族，其类别构成 [15]中式(1.1)的 Ω 。嵌套于结构族之内的是由附加条件式(3.2)定义的历史存储库族，该条件再现了观测到的绿色电力缺口。从概念上讲，存储库族是由一组线性丢番图方程确定的离散对象。在 [15], 中式(2.21)的$\widehat{}$单纯形上，我们类似地通过包含归一化在内的d个 $= 3$ 与D个 $= 4$ 线性独立条件，分别定义分布的结构 $\mathcal{L}$ 族与嵌套的历史$\mathcal{A} \equiv \mathcal{L}_{\text{obs}}$ 族，即 [15] 中的式(5.2)与式(5.55)。

给定一个假设反事实，我们将生成规模为 N 的替代存储库（经粗粒化后归为存储库族，即备择假设数据集族），视为属于多项抽样普适类。这种一般抽样从假设分布 $\mathfrak{v} = \mathfrak{h}$ 中进行，该分布属于第 $\mathcal{H}$ 节组合假设所满足的所有分布构成的假设线性族3。1。我们将一般抽样限定在

存储库结构族中，其中所有库的总容量和水力发电容量均相同。在样本量 N 较大时，根据p，观测到的绿色电力赤字的内在5.14式和5.69式， [15], 可近似为 1 减去自由度为 $\mathrm{k} = D - d = 1$、检验统计量为以下数值的 χ^2_k分布的累积分布函数：

$$t = 2N \min_{\mathfrak{p} \in \mathcal{A}} \mathrm{D}(\mathfrak{p} \parallel \hat{\mathfrak{q}}) \quad \text{with} \quad \hat{\mathfrak{q}} = \operatorname*{argmin}_{\mathfrak{p} \in \widehat{\mathcal{L}}} \mathrm{D}(\mathfrak{p} \parallel \mathfrak{h}) \,. \tag{A.1}$$

在我们当前的应用中， $\mathcal{H} \subset \mathcal{L}$ 因此$\hat{\mathfrak{q}} = \mathfrak{h}$ 显然成立，因为 $\mathfrak{h}$ 本身就是 $\mathcal{L}$ 中一个可行的成员，因此也是后者信息散度的全局最小值。根据构造，历史存储库属于历史存储库族，因此 $\mathfrak{f} \in \mathcal{A}$。

玻尔兹曼假设分布的推导。在第3.2节中，我们阐述了为何要从假设线性族 $\mathcal{H}$中选取一个具有特定性质的成员，即玻尔兹曼假设分布；它在操作层面和治理场景中都具有理想的意义。现在，我们将从最大概率条件推导出玻尔兹曼假设分布。在玻尔兹曼假设反事实下，该条件使历史存储库成为历史存储库族内条件抽样的众数。

我们首先将证明，在单纯形上，并在 Csiszár 基于大N 一般抽样处理所形成的信息几何表述框架内，最大概率条件唯一确定玻尔兹曼假设分布。随后我们将证明，即使在有限N的情况下，玻尔兹曼假设分布赋予历史存储库的多项概率，也高于历史族中所有其他项目库。

考虑历史项目库，其组成由经验分布 $\mathfrak{f} \in \mathcal{A}$ 描述。为使p值具有可解释性而提出的认识论论证，选择假设分布 $\mathfrak{h} \in \mathcal{H}$ ，使得 $\mathfrak{f}$ 描述历史存储库族中概率最高的项目库的组成。对历史存储库族内抽样的多项概率进行标准的玻尔兹曼–萨诺夫大N 展开并截断后，这一要求得出

$$\mathfrak{f} \overset{!}{=} \operatorname*{argmin}_{\mathfrak{p} \in \mathcal{A}} \mathrm{D}(\mathfrak{p} \parallel \hat{\mathfrak{q}}) = \operatorname*{argmin}_{\mathfrak{p} \in \mathcal{A}} \mathrm{D}(\mathfrak{p} \parallel \mathfrak{h}) \,. \tag{A.2}$$

第二个等式成立，是因为我们的结构条件将Y 的相关广义矩固定为其经验估计。进而，这将检验统计量式（A.1）简化为$t = 2N\mathrm{D}(\mathfrak{f} \parallel \mathfrak{h})$。

在以下推导中，我们仅考虑组合假设未将任何观测项目概况的权重降为零的情况。因此，相关的分布线性族具有非空内部，且下文考虑的所有 I-投影都是内部点。利用上文式（A.2）中的条件，我们便可证明

$$\mathfrak{h} = \operatorname*{argmin}_{\mathfrak{p} \in \mathcal{H}} \mathrm{D}(\mathfrak{p} \parallel \mathfrak{f}) \,. \tag{A.3}$$

因此，在一般抽样的情况下，具有认识论意义的假设分布

在线性条件下，可以理解为通过将经验分布经由I-投影到假设线性族上，实现最小适应原则。

首先，我们注意到，假设分布通常不会分解，例如$h(y,x,s) = h(y,s)h(x)$，即使在式（3.3）这类简单线性条件下也是如此，因为我们处理的是观测到的（而非可观测的）项目概况。对于其余形式化断言，按索引 $i \in \{1,\dots |\Omega|\}$ 至 $|\Omega| \le N$ 对观测到的概况进行编号会很方便。由于聚合条件关于概率是线性的，因此可以通过线性系统（式（2.22）中的 [15]）加以概括。

$$\mathbf{G}\,\mathfrak{p} = \boldsymbol{\mu}\,. \tag{A.4}$$

这里，$\boldsymbol{\mu}$ 是一个取实数值的广义矩列向量。在我们的应用中，对于结构族，$\hat{}\boldsymbol{\mu}$ =(1, $\bar{}$Y, $\bar{}$Y $_{\text{hydro}})^T$ ；对于备择假设，则为$\boldsymbol{\mu}_A$ = (1, $\bar{}$Y, $\bar{}$Y$_{\text{hydro}}$, $\bar{}$Y$_{\text{green}}$ − $\bar{}$Y$_{\text{brown}})^T$ ；对于零假设族，则为 $\boldsymbol{\eta} = (1, \bar{}\text{Y}, \bar{}\text{Y}_{\text{hydro}}, 0)^T$ 。在简写记法中，$\bar{}$Y$\bar{}$表示总铭牌容量（Y 的样本均值），而 Y_s 表示具有给定政策标签 s 的剖面上的总量（细粒度的 ˜s 也相应如此）。适当系数矩阵 G 的矩阵元素 $g_{a,i}$ （矩阵有 d 行或 D 行，以及 $|\Omega|$ 列）可根据相应条件读出（对项目概况编号后即可）；在我们的应用中，这些元素由铭牌容量 Y 以及政策 S 和˜细粒度 S 可观测量的指示函数给出。

在引入必要的记号和线性代数形式后，我们现在证明，公式（A.3）足以推出公式（A.2），并且在大N 信息几何表述中也成为必要条件。充分性直接源于公式（I-投影（A.3）的指数形式；对于某个拉格朗日乘子向量 $\widehat{\boldsymbol{\theta}} \in \mathbb{R}^D$，该式可写为

$$h_i = f_i \,\exp\left\{\sum_{a=0}^{D-1} \widehat{\theta}_a\, g_{a,i}\right\}\,. \tag{A.5}$$

对于可行的联合假设，这种 I-投影的参数化始终存在，并且在线性系统 $\mathbf{G}\,\mathfrak{h} = \boldsymbol{\eta}$ 的重新参数化下保持唯一性；这种重新参数化会逆变地变换对偶参数，从而使概率 h_i 保持不变。将公式（A.5）的形式代入公式（A.2）中的首项信息散度，得到

$$\mathrm{D}(\mathfrak{p} \parallel \mathfrak{h}) = \mathrm{D}(\mathfrak{p} \parallel \mathfrak{f}) - \widehat{\boldsymbol{\theta}} \cdot \boldsymbol{\mu}_A \quad \text{for} \quad \mathfrak{p} \in \mathcal{A}\,, \tag{A.6}$$

因为历史族中的所有分布 $\mathfrak{p}$ 都会再现实证估计值 $\boldsymbol{\mu}_A$ （广义矩）。由于这些经验估计对于给定的历史存储库而言是常数，并且拉格朗日乘子由组合假设确定，所以公式（A.6）在 $\mathcal{A}$ 上达到最小值，当且仅当从 $\mathfrak{f}$ 出发的信息散度达到最小值。在历史存储库族上，后者显然在 $\mathfrak{p} = \mathfrak{f}$处达到最小值，由此证明了充分性。

反之，假设已根据式（ $\tilde{\mathfrak{h}} \in \mathcal{H}$ A.2）中的最大概率原理找到了。在历史存储库族条件下，从 $\tilde{\mathfrak{h}}$ 进行抽样的模式（当N足够大时）由I-投影 $\tilde{\mathfrak{h}}$ 到 $\mathcal{A}$来刻画。改写

其指数形式可得：

$$\tilde{h}_i = f_i \exp\left\{ -\sum_{a=0}^{D-1} \widehat{\lambda}_a \, g_{a,i} \right\} , \tag{A.7}$$

对于某个 $\widehat{\boldsymbol{\lambda}} \in \mathbb{R}^D$。由此，我们认识到，这个 $\tilde{\mathfrak{h}}$是指数族$\mathcal{E}$ 的成员，其

$$p_i(\boldsymbol{\theta}) = f_i \exp\left(\sum_{a=0}^{D-1} \theta_a \, g_{a,i} \right) \quad \text{for} \quad \boldsymbol{\theta} \in \mathbb{R}^D , \tag{A.8}$$

围绕经验分布生成 $\mathfrak{f}$，其系数矩阵为 $\mathbf{G}$。因此， $\tilde{\mathfrak{h}} \in \mathcal{H} \cap \mathcal{E}$。根据线性族与指数族之间的对偶性， [30], 该交集是唯一的，并且与I-投影一致，即将 $\mathfrak{f}$ 投影到 $\mathcal{H}$上，因此 $\tilde{\mathfrak{h}} = \mathfrak{h}$。

至关重要的是，式（A.3）中针对玻尔兹曼假设分布的规定，继续保证原始历史存储库在有限 N 时，在历史存储库族中仍被赋予最大概率。在多项抽样的普适类中，任何过程的关于 N 的首项抽样概率都必须由多项分布决定（见 [15]的附录 A），因此只需考察

$$N\mathfrak{p} \in \mathbb{N}_0^{|\Omega|} , \; \mathfrak{p} \in \mathcal{A} \; : \quad \log \mathrm{mult}(N\mathfrak{p}; \mathfrak{h}) = \log \mathrm{mult}(N\mathfrak{p}; \mathfrak{f}) + N \, \widehat{\boldsymbol{\theta}} \cdot \boldsymbol{\mu}_A \quad , \tag{A.9}$$

代入式（A.5）中的指数形式后。后一项在整个备择假设族中保持不变，因为它只取决于原始组成 $N\mathfrak{f}$、系数矩阵 $\mathbf{G}$ 以及假设的广义矩 $\boldsymbol{\eta}$。

因此，对于 $\mathfrak{h}$ ，在历史存储库族中对多项概率进行精确最大化，等价于对 $\mathfrak{f}$进行相应的最大化。众所周知，参数（内禀概率）为f_i 的多项分布在计数向量$N\mathfrak{f}$处达到全局最大值，该向量包含历史剖面频率。由于历史存储库族（与其他任何仓库族一样）是大小为N 的可能项目库的子集，并且按定义包含历史存储库本身，因此计数向量$N\mathfrak{f}$必然使历史存储库族上的多项概率最大化。根据前面对式（A.9）的分解，它也因此使来自 $\mathfrak{h}$的抽样概率最大化。这个有限N 充分性论证适用于任意 $|\Omega| \leq N$。

反事实广义矩对历史基线的可归因性。遵循以下信息论推导 $\mathfrak{h}$ 并确认其在认识论上令人满意的性质后，我们现在揭示玻尔兹曼假设分布的一项更严格性质，使其在 $\mathcal{H}$的各个成员中独一无二。在玻尔兹曼假设反事实中，凡是第3.1 节的组合假设未予指定的广义矩，其下游取值都可通过条件抽样下的可归因性唯一确定。具体而言，条件抽样估计值[8] 仅当支持历史绿色电力

[8] 显然，无条件样本估计会围绕假设分布所规定的反事实值波动。

赤字的替代存储库是在玻尔兹曼假设反事实下生成时，该广义矩的估计值才会围绕其历史值波动。

为展示这一点，当 N 和项目概况数量都超过强加条件的数量时，我们考虑一个函数 $\phi : \Omega \to \mathbb{R}$该函数规定了一个广义矩，而该广义矩最终可能会纳入某个下游指标的定义中。我们关注的是其他方面任意、且不能由Y矩推出的广义矩，这些Y矩会纳入结构条件和强制条件。在结构仓库族中，我们随后构造两个备择假设仓库族，它们都符合结构条件。历史仓库族通过指定观测到的功率亏缺（式（3.2））来定义，与之前相同。嵌套仓库族还支持广义矩的历史估计值，

$$\sum_{i=1} p_i \phi_i \stackrel{!}{=} \sum_{i=1} f_i \phi_i \,, \tag{A.10}$$

因此总计由五个线性独立条件定义，其中包括归一化。

给定某个假设分布 $\tilde{\mathfrak{h}} \in \mathcal{H}$，我们进行嵌套假设检验，以评估在给定观测到的绿色电力缺口的情况下，广义矩的历史估计值在统计上是否合理。具体而言，我们比较嵌套族与历史族的抽样概率之比。比值接近 1，意味着式（A.10）中的附加条件几乎不会损失概率质量，因此，与 ϕ 的历史样本均值之间的偏差可归因于历史项目库族内的非系统性波动。因此，我们会得出结论：在指定的结构条件下，了解观测到的绿色电力缺口，在统计上意味着附加广义矩的历史估计值。

对该比值进行鞍点近似，有助于在给定历史绿色电力缺口的情况下，计算观测到历史估计值或广义矩中任何至少同样不可能值的内在p值。近似 χ^2_1 的检验统计量为

$$t = 2N \min_{\mathfrak{p} \in \mathcal{A}_{\text{nested}}} \mathrm{D}(\mathfrak{p} \,\|\, \mathfrak{q}) \quad \text{with} \quad \mathfrak{q} = \operatorname*{argmin}_{\mathfrak{p} \in \mathcal{A}} \mathrm{D}(\mathfrak{p} \,\|\, \tilde{\mathfrak{h}}) \,, \tag{A.11}$$

此处，历史族 $\mathcal{A}$ 现在在单纯形上发挥环境线性族的作用，参见式（A.1）。如上所示，$\mathfrak{q}$作为投影到历史线性族上的 I-投影，除非我们选择玻尔兹曼假设分布，否则不会与 $\mathfrak{f}$一致。类似地，$\mathfrak{p} = \mathfrak{q}$不必满足式（A.10），因此环境 I-投影通常不会属于嵌套族 $\mathcal{A}_{\text{nested}} \subset \mathcal{A}$。

要求由任何此类广义矩所捕捉的 $\tilde{\mathfrak{h}}$-反事实结构，在给定历史绿色电力缺口的情况下，完全归因于历史存储库，这意味着式（$A11$）中的内在 p-值为 1。通俗地说，我们要求：在所选假设反事实的全集中，知道绿色电力缺口就意味着知道其他任一广义矩的历史估计值。玻尔兹曼假设反事实自动满足这一要求，因为 $\mathfrak{q} = \mathfrak{f} \in \mathcal{A}_{\text{nested}}$ 对于 $\tilde{\mathfrak{h}} = \mathfrak{h}$ 成立，因此$t = 0$。

现在我们选择另一个 $\tilde{\mathfrak{h}} \neq \mathfrak{h}$，它可得出 $\mathfrak{q} \neq \mathfrak{f}$。由于 ϕ 是任意的，且 Ω 是有限的，只要 $\mathfrak{q} \neq \mathfrak{f}$，我们总能选择一个广义矩，使得 I-投影到

环境历史族不满足式（A.10），因此 $\mathfrak{q} \notin \mathcal{A}_{\text{nested}}$。进而，这会得到一个非平凡的 I-投影到嵌套族，这意味着式（$t > 0$ （A.11）中的内容。总之，对于任何单独考察的广义矩，选择玻尔兹曼假设分布是确保反事实估计可归因于历史基准的必要且充分条件。给定观测到的绿色电力缺口后，该论证可推广至多个广义矩的嵌套检验。

内禀检验统计量对经验分布的显式依赖。
方程（A.2）还带来进一步的解释性推论，表明内禀检验统计量具有一种理想性质。另一个通过某种其他机制选出的假设分布 $\tilde{\mathfrak{h}}$，在对存储库族上的抽样概率进行粗粒化后，通常会得到一个用于内禀p值的检验统计量，其形式为$t' = 2N\mathrm{D}(\tilde{\mathfrak{q}} \parallel \tilde{\mathfrak{h}})$，其中 $\tilde{\mathfrak{q}} \neq \mathfrak{f}$ 是 $\tilde{\mathfrak{h}}$ 在历史分布族上的 I-投影。

现在假设，针对 $\tilde{\mathfrak{h}}$ 的选择机制无法察觉经验分布 $\mathfrak{f}$ 的变化，而这种变化除了 $\tilde{\mathfrak{h}}$ 之外，不会改变定义结构族和假设族的广义矩。对于政策情境中考虑的欠充分组合假设，这类方向总是存在。因此，I-投影 $\tilde{\mathfrak{q}}$ 及由此得到的统计量t' 保持不变。换言之，这类变化中所包含的经验信息已经丢失，实际上被一种依赖于“模型”的规定所取代；该规定通过针对 $\tilde{\mathfrak{h}}$ 的选择机制来定义 $\tilde{\mathfrak{q}}$。

相比之下，式（A.2）确保内在p值的检验统计量简化为$t = 2N\mathrm{D}(\mathfrak{f} \parallel \mathfrak{h})$，因此，历史存储库的变化会直接进入检验统计量，而不会在粗粒化后才被吸收。从这个意义上说，最大概率条件可防止内在p值获得间接且依赖于“模型”的检验统计量。在我们的玻尔兹曼构造中，如果要调整项目库组成的某些方面，可以将其直接指定为额外的结构条件；这些条件对应的广义矩不一定要由其经验估计给出。

与以模型为中心的文献的关系。此处推导的假设分布在式（A.5）中的指数倾斜形式，在形式上与计量经济学中的倾斜构造相似。在这类文献中，为了在假定的统计模型内进行估计，人们会最小化信息散度（以及其他散度），以构造满足矩约束的概率权重；这一做法符合概率论的公理化表述。参见例如 [31–33]。

相比之下，在我们的非公理化构造中，指数形式是由大N 的一般抽样与最大概率条件共同产生的。该条件要求历史存储库成为历史存储库族中概率最高的成员，从而选出玻尔兹曼假设分布。随后，玻尔兹曼假设分布被用于对存储库族上的抽样概率进行粗粒化，由此导出可解释的内禀p值。

历史存储库的典型性。有趣的是，式（A.2）中的最大概率原理确定了（在 N 的领先阶上）对历史存储库族进行抽样时的期望频率，至少[9] 只要 $|\Omega| \ll N$ 属于多项抽样的普适类。将通常的熵集中论证加以推广，[34], 我们可以用拉普拉斯近似估算

$$\sum_{\mathbf{n}\in\mathbb{N}_0^{|\Omega|}} \mathbf{n}\,\Pr(\mathbf{n}|\mathrm{A};\mathfrak{h}) = N\mathfrak{f} + \mathcal{O}(1)\ , \tag{A.12}$$

通过在连续统极限下，对 $\mathfrak{h}$ 投影到 $\mathcal{A}$ 的I-投影附近的抽样概率密度进行高斯化，参见例如 [15] 的第 2 节。求和遍历历史存储库族 A 中的所有库，即遍历所有能够再现观测到的重缩放广义矩 $N\boldsymbol{\mu}_A = \mathbf{G}\mathbf{n}$的计数向量$\mathbf{n}$ 。

因此，认识论论证表明，在依据玻尔兹曼假设分布进行条件抽样时，历史项目库不仅是最有可能的，而且在期望意义上也是典型的 。换言之，在玻尔兹曼假设反事实下，历史相对频率使抽样概率最大；从历史存储库族中抽样备择项目库时，它们与相应的期望相对频率之差为 $1/N$ 阶修正。

A.3 更下游分析

我们对第4.2节的下游分析进行了补充，进一步分析为绿色转型构建的玻尔兹曼假设反事实中可能值得关注的其他维度。

跨铭牌容量的重加权。为更细致地考察强制转型所依据的项目概况重加权，我们考虑铭牌容量Y 与政策可观测量 S的边缘分布。在图 8 中，我们将历史边缘分布 $f(y,s)$ 与零功率差处玻尔兹曼假设分布 $h_{\text{null}}(y,s)$ 的相应边缘分布进行比较，此时绿色与棕色电力总量恰好平衡。

图 8 所示的重加权揭示了比两种技术标签出现频率的简单变化更深层次的重构。具体而言，铭牌容量较高的绿色项目概况获得了更高权重，而绿色项目概况在较低Y 端的权重则降低。同时，棕色项目概况的权重也有所降低，且这种降幅在较高Y 端变得更加明显。固定的聚合水力发电容量对最高Y 端的绿色项目概况形成约束，进一步限制了通过此类水力发电项目概况获取额外绿色电力的程度。因此，最终的重加权方案以非平凡的方式取决于项目概况的技术标签和铭牌容量。[10]

[9] 由于 $\delta \in \ker\mathbf{G} \Rightarrow -\delta \in \ker\mathbf{G}$，围绕 $N\mathfrak{f}$ 的涨落之间的高阶抵消，即使在 $|\Omega| = \mathcal{O}(N)$ 下“朴素”的连续统极限下的高斯化失效时，仍可确保估计式（A.12）的稳定性。[10] 特别地，我们的组合假设并不采用试探形式 $h_{\text{null}}(y,s) = f(y\,|\,s)h_{\text{null}}(s)$ 来约束

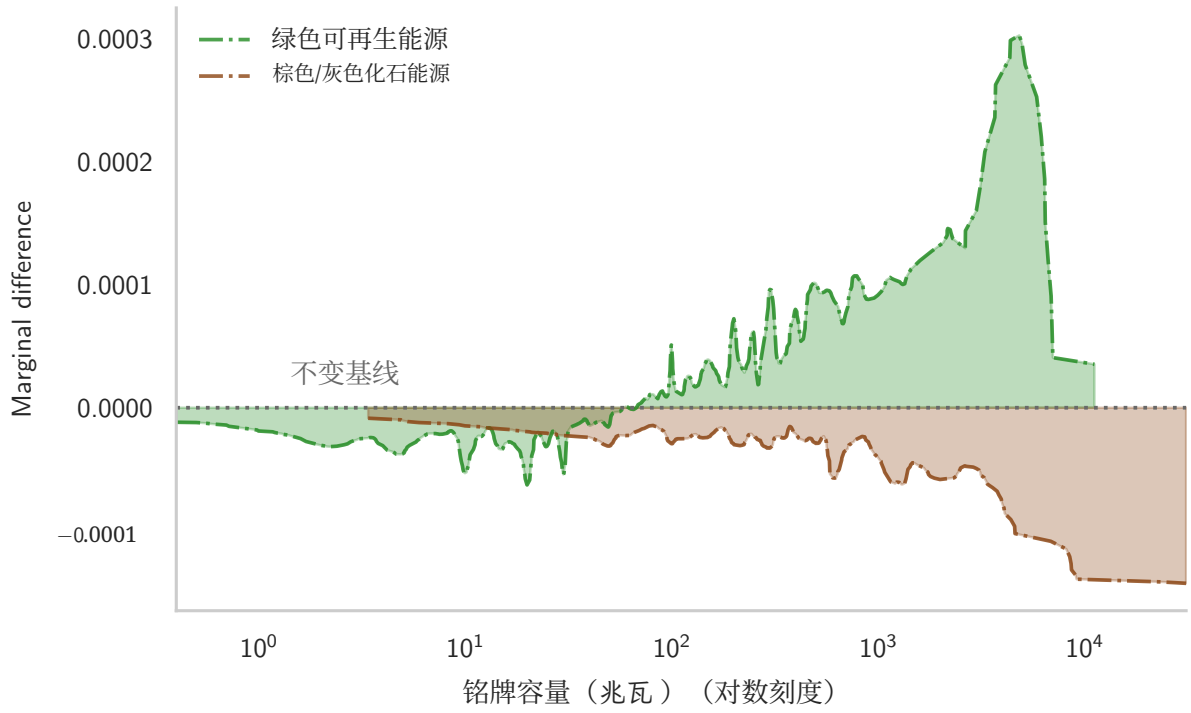


图 8：**零假设下铭牌容量的重加权**。曲线分别针对每种技术标签，展示了零功率差反事实假设下以及历史存储库中的Y 和S 边缘分布之差。为便于观察，曲线经过宽度为 $\sigma = 5$ 的高斯滤波器平滑处理，以揭示在绿棕功率精确平衡时，反事实重加权如何随观测到的铭牌容量全范围变化。平滑处理前，零差基线上方增加的概率质量与其下方移除的概率质量恰好平衡，从而保持总概率归一化。

如图所示，满足少量聚合条件会导致概率质量在各个单个项目概况之间发生非平凡的重新分配；这种分配既无法从聚合条件本身直接看出，也难以用闭式表达。一般而言，重加权的影响会通过经验分布中编码的非平凡关联，以非平凡的方式传播至下游分析所考虑的所有维度，即使组合假设仅约束Y 和S 可观测量。

私人所有权参与的稳定性。在分析了Y 与 X 在 S条件下的反事实联合统计后，我们现在转向次级可观测量 Z_1，，它对应于 `PercentPrivate` 属性，记录了私人出资方拥有的项目公司股份百分比 [18]。同样，我们按照政策可观测量 S 对下游分析进行分层，即在假设反事实中，同时考察条件边际概率 $h(z_1|s)$ 和边际概率 $h(z_1)$ 。

在图9中，我们考察了不同目标功率差下预期的私人所有权份额。按技术标签分层的条件均值显示，在整个绿色电力主导转型过程中，私人所有权

给定技术标签的铭牌容量条件分布等于其历史对应值，即我们不考虑更严格的一组条件 $h_{\text{null}}(y \mid s) = f(y \mid s)$ 对所有观测到的元组 (y, s)均成立。

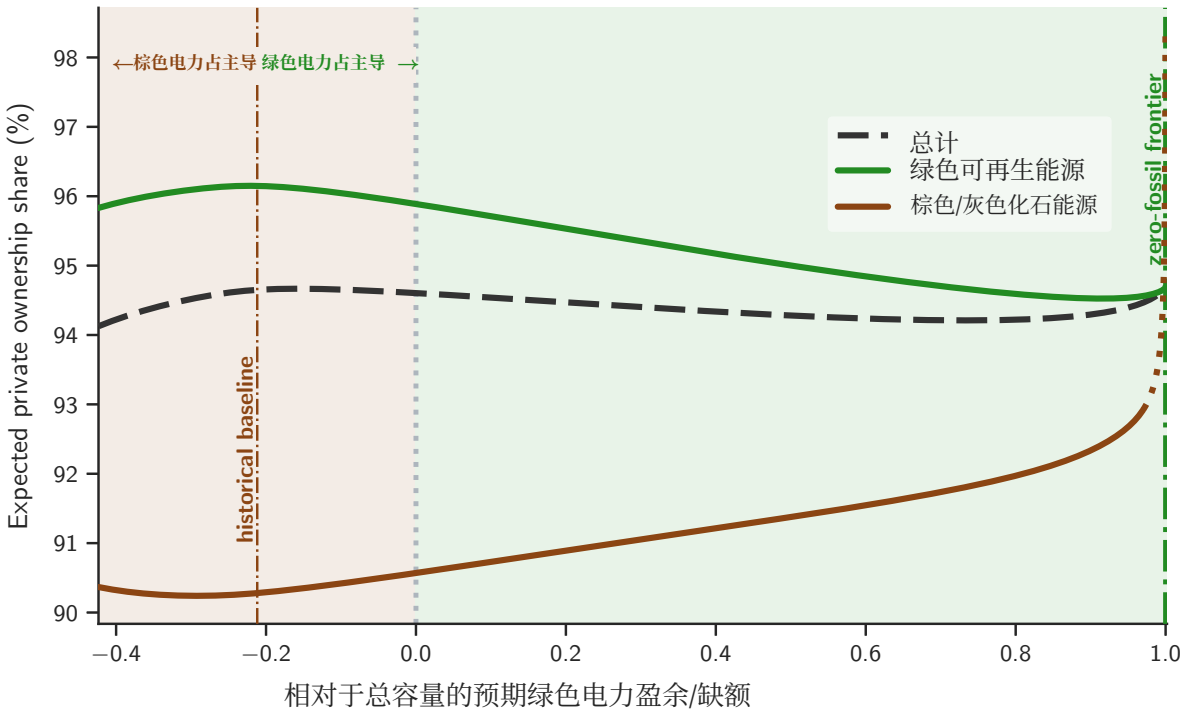


图9：**预期私人所有权份额**。按技术标签分层的私人所有权份额条件均值，以及玻尔兹曼假设反事实中的总体均值。

行业参与度相当稳定。在历史基线附近，绿色项目中的预期私人所有权份额略高于 96%。随着规定的电力差异增大并朝绿色主导方向发展，这一份额逐渐下降，但即使在零化石能源边界处仍高于 94%。

棕色项目的相应份额始终系统性偏低，在棕色项目仍普遍存在的地区内，波动于相对狭窄的90%至93%区间。随着接近零化石能源边界，表面上的急剧上升（由棕色虚线段标示）并不反映玻尔兹曼重新加权的不稳定性。相反，它源于棕色项目流行度趋于消失，$h(\texttt{brown}) \to 0$。随着棕色项目的流行度逐渐变小，相应的条件均值变得愈发敏感，并最终在零化石能源极限处无法定义。

总体预期私人所有权份额的变化也十分有限，始终高于94%。回想一下，总体均值是两个条件均值的按流行度加权的平均值，而带标签的流行度 $h(s)$ 本身沿目标功率差变化（参见图3）。因此，在假设反事实中，总体私人所有权参与对目标环境要求大体不敏感。特别是，当目标功率差趋向绿色主导时，该反事实并未表明公共所有权会大幅取代私人所有权。

项目库的轻微时间结构调整。在考察绿色转型对财政和所有权的影响后，我们现在转向历史项目库的时间构成，该项目库涵盖了40年间的发电项目

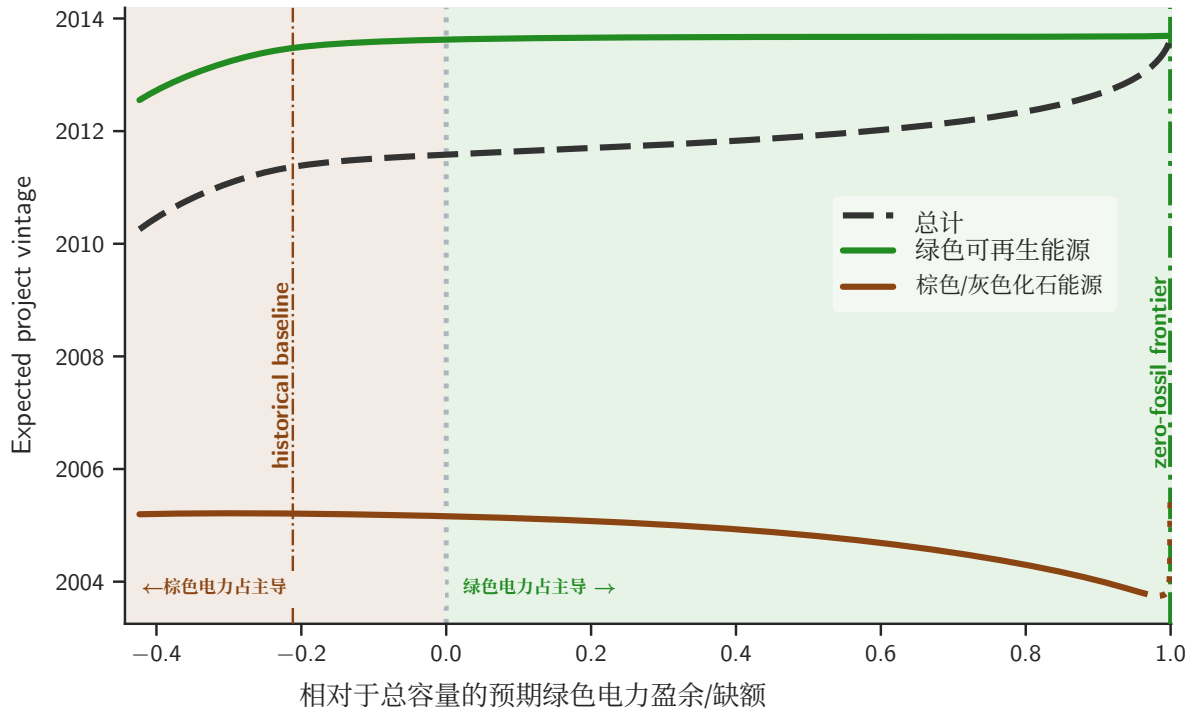


图 10：**预期项目年份**。按标记技术分层的项目年份条件均值，以及玻尔兹曼假设反事实中的总体均值。

自 1984年 至 2024年。为分析这一构成如何调整，以在玻尔兹曼假设反事实中实现绿色强制要求，我们在图10中考察不同目标电力差异下的预期项目年份。表示投资年份的可观测量T属于第2.3节介绍的用于记账的核心可观测量经济组。同样，我们按照政策可观测量S对下游分析进行分层，即在假设反事实中，同时考察条件边际概率$h(t|s)$ 和边际概率$h(t)$ 。

按技术标记分类的项目年份条件均值在整个转型过程中相对稳定。绿色轨迹从历史基线开始逐步上升，并在略低于 2014 年处趋于稳定，此时已处于棕色主导区域；棕色轨迹则仅从略高于 2005 年小幅下降至约 2004 年。因此，总体均值年份从赤字极端情况下的约 2010 年，变为历史基线下的 2011 年；随后，随着棕色项目占比逐渐消失，总体均值逐渐接近绿色条件均值。随着棕色项目占比趋于零化石能源边界而消失，棕色条件均值会出现已知的数值不稳定性，图中采用与上文相同的点线表示法加以展示。

地理分配。我们通过财政可观测量 X、政策可观测量 S和次级可观测量 Z_2 所对应的 `Region` 属性 [18]的联合边际分布，考察构建的反事实中的空间维度，以此完成下游分析。图 11

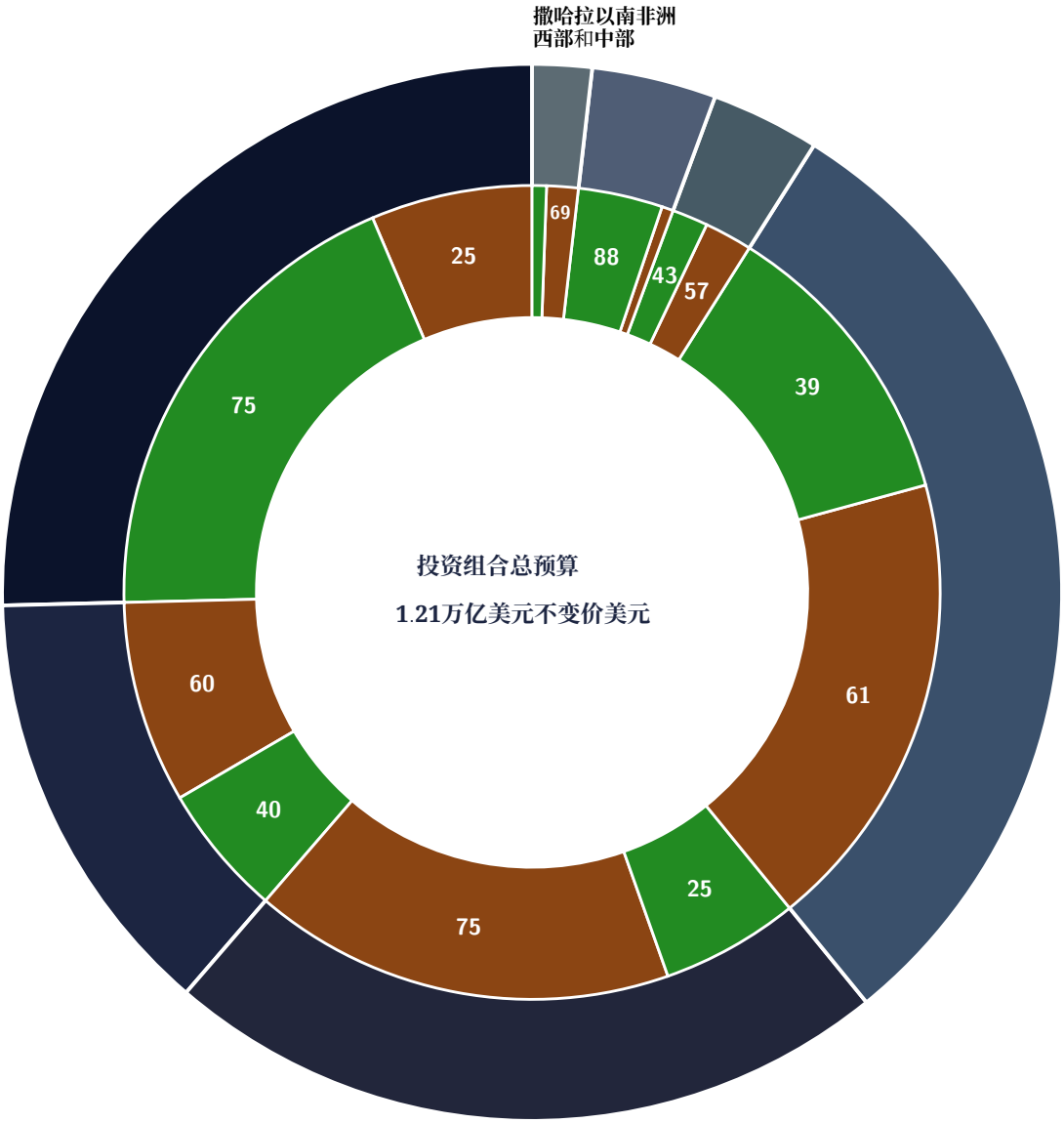


图11：**历史存储库中的区域投资份额**。同心圆饼图显示了总投资在七个宏观区域之间的分配情况，其中每个区域的份额进一步划分为绿色项目和棕色项目。内环中显示的百分比四舍五入到最接近的整数；为便于视觉识别，省略了%符号。

和12。外环显示各地区的投资份额，

$$\sum_{x} xp(x, z_2) \Big/ \sum_{x} xp(x) \quad \text{for} \quad z_2 \in \{\texttt{Latin America \& the Caribbean}, \ldots\}\,, \tag{A.13}$$

而内环则将每个区域的份额划分为绿色与棕色投资，

$$\sum_{x} xp(x, z_2, s) \Big/ \sum_{x} xp(x, z_2)\,. \tag{A.14}$$

图 11 展示了世界银行数据库中总投资在七个地理区域间的历史分配情况（$\mathfrak{p} = \mathfrak{f}$），并按区域明确区分预算中分配给绿色与棕色项目概况的比例。

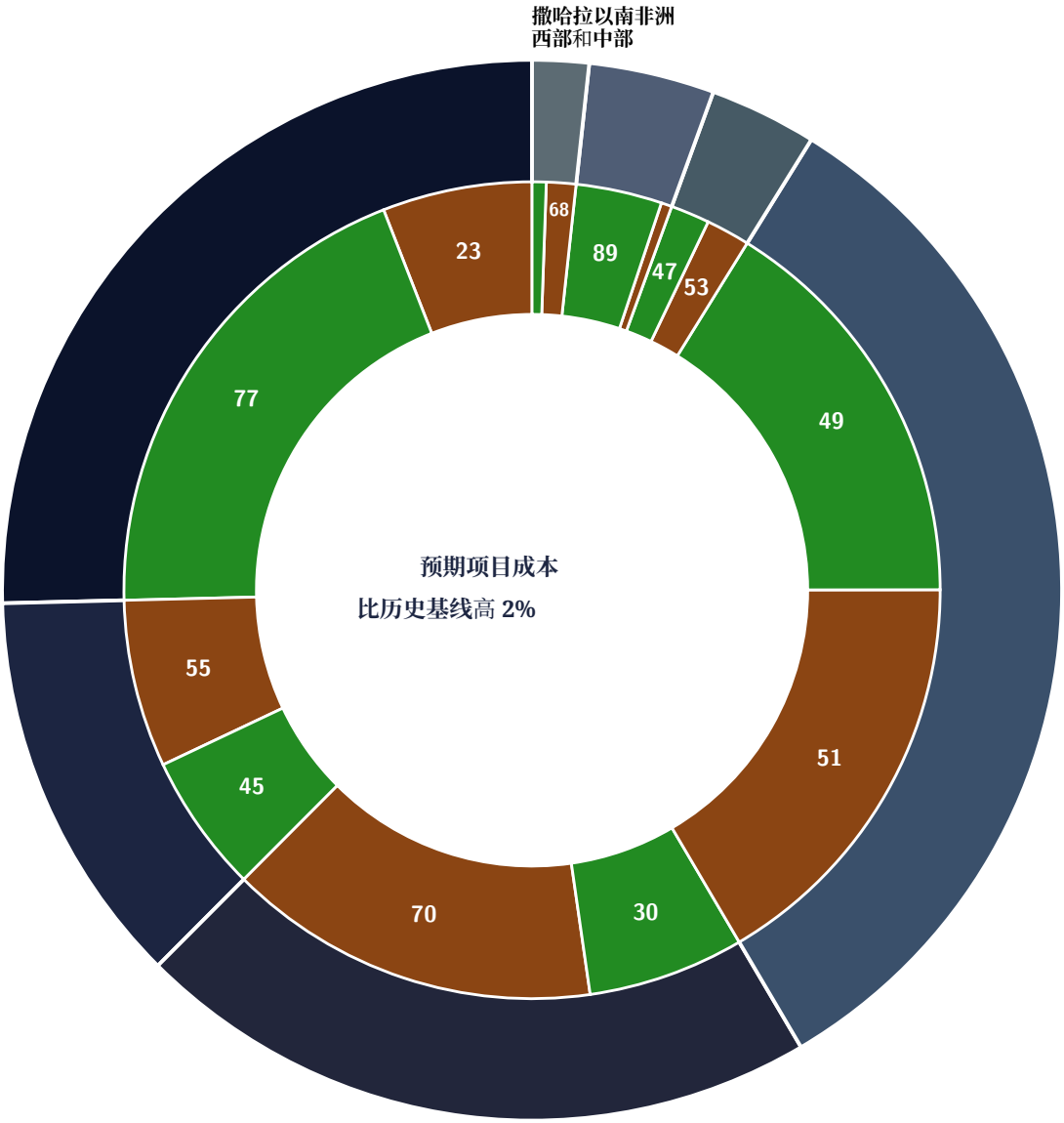


图12　　：　**绿色与棕色电力平价下的区域投资份额，** $\eta = 0$ **见式** (3.1　　).

图表明，世界银行数据库中的地理区域，明确细分了各区域预算中分配给绿色与棕色项目概况的比例。图 12 展示了零电力差下转型机制的玻尔兹曼假设反事实（$\mathfrak{p} = \mathfrak{h}_{\text{null}}$）中的空间构成。直接比较这两种空间构成，会得出一个如今已在意料之中的结果。分配给各宏观区域的投资份额几乎没有变化，调整幅度微乎其微，因为绿色强制要求主要通过各宏观区域内部将投资从棕色转向绿色来实现。