

教师聘用网络中的系统不平等和等级制度

Clauset, A., Arbesman, S., & Larremore, D. B.

左玥、尹凤杰、汪茹译，沈文钦校

原文载于：

Clauset, A., Arbesman, S., & Larremore, D. B. (2015). Systematic inequality and hierarchy in faculty hiring networks. *Science advances*, 1(1), e1400005.

教师就业市场在形塑科学家和科研机构的研究重点、教育成果和职业轨迹方面发挥着根本作用。然而，**当前缺少将教师聘用作为一个系统的量化理解**。为了解哪些机构聘用了哪些机构的毕业生作为教师，我们利用一种简单的技术来提取最能解释所观察到的教师聘用网络的机构声望排名，展示并分析了**三个不同学科**近 **19,000 名** 普通教师的综合职业数据。我们发现不同学科中教师的聘用都遵循着共同且分化明显的等级结构，这反映出严重的社会不平等。此外，仅靠博士学位声望就能比美国新闻与世界报道的排名更好地预测最终职位，女性通常比男性的地位更差。声望的提高会使某一机构的博士毕业生获得更多、更好的教师职位，并增强其在本学科领域的影响力。这些分析结果提高了我们对学术界声望影响力进行量化的能力，使我们对学术系统有了新的认识。

教师招聘是学科的一个普遍特征，哪些机构聘用了哪些院校的毕业生，其结果几乎形塑着学术生活的各个方面，包括学术生产力，研究重点，资源分配，教育成果，以及学者个人的职业轨迹（1-4）。尽管教师聘用的功能如此重要，但对于各学科中教师聘用的通用模式和效率，我们仍然缺乏清晰和系统的理解。

从制度角度看，教师聘用是一种隐性的评估：当一所机构 **u** 将另一所机构 **v** 的毕业生聘为教师时，**u** 便对 **v** 的教学和科研质量作出了正面评价。同样地，当个体接受了 **u** 提供的职位，他也就对 **u** 的质量作出了正面评价。作为这种互评的集合，一个学科的教师聘用网络（图 1）便代表了其自身的教学和科研成果的总体评估情况（5）。不同院校在生产教师方面的能力是不

同的，在其他较成功的院校获得更多职位也就意味着该院校的成果得到了更高的总体评价。

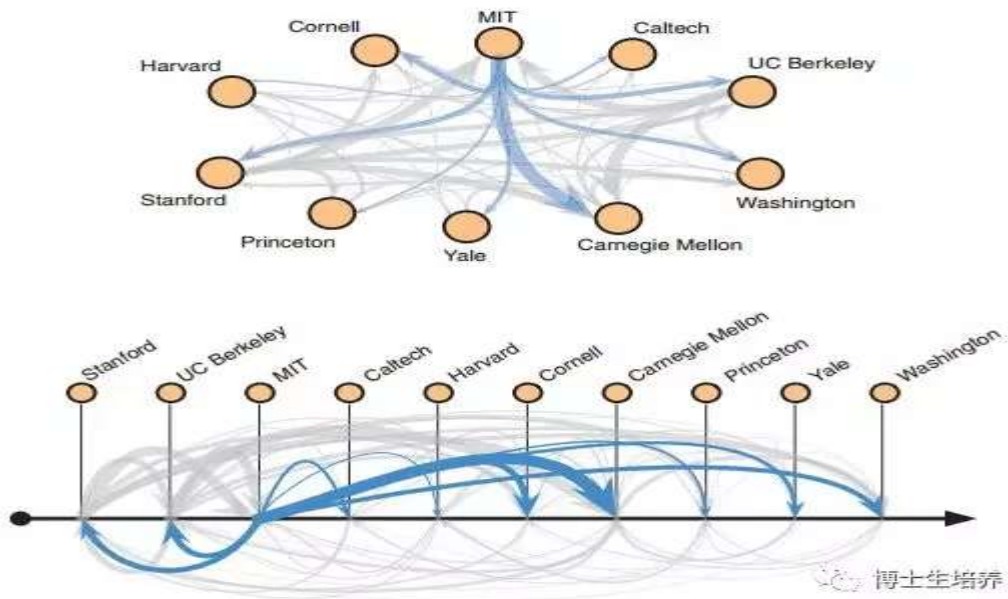


图 1 职员聘用网络声望等级制度

在这种竞争中，不同的成功率是社会等级的标志，可能来自物质上的优势或社会声望机制（6）。在学术机构中，物质上的优势可能会被忽略，社会威望成为主要因素。声望较低者可能会努力模仿声望较高者的成功行为，以求提高自己的声望（7,8）。在这种情况下，教师聘用中的声望因素是一个可操作变量，它包含了学术优势和非能力性因素（如社会地位或地理位置）的差异。如果社会地位、地理位置等因素不相关，那么声望就相当于才能（merit）。更现实的是，非才能因素确实存在影响，而且其重要性越大，声望与才能之间的相关性就越小。

机构声望依赖于制度之间的相互作用、主观评价等其他因素，因此测量机构声望是比较复杂的。如美国新闻、世界报道和国家研究委员会（NRC）的权威排名（9）等经典的方法往往独立地量化机构，忽略了诸如联合计划、研究合作、研究生招生或教师招聘等互动的影响。这样的排名也因为强调声誉、财富和“选择性”等投入而非教育产出而受到了广泛的批评（10,11）。相反，教师聘用网络同时反映了交互和专家对于成果的评估，这使得通过一种有效的、量化的途径来描述声望的影响、确定大规模的招聘模式并揭示才能和地位的相对作用成为可能。

在这里，我们人工收集了 461 个北美学系或学院级别的学术单位的教师数据，以分析教师聘用网络的结构，这些数据涉及计算机科学、商学和历史专业，包含近 19,000 位终身轨或获得终身教职的教师（见补充材料和表格 S1）。这些学科代表了截然不同的学术传统，为描绘学术界教师职位的一般模式提供了广泛的基础。我们样本中的机构是从每个学科的博士授予单位综合名单中选出的。一位教师必须在我们的某一样本高校中获得博士学位，并在某一样本高校中任职，才能进入我们的数据库。在抽样的教师中，有 86% 符合这些标准，意味着这些机构中存在一个几乎封闭的博士生生态系统。

对于这些数据，我们应用了一种新的基于网络的技术来提取最能解释观察到的聘任决定的声望等级。在不同学科当中，我们表明教师聘用遵循着一个共同的层次分明的等级结构，它反映出机构中潜在的社会不平等现象。此外，我们也表明（1）仅靠博士学位声望就能比美国新闻、世界报道和国家研究委员会的排名**更好地预测**最终职位，（2）来自同一机构的博士毕业生中，女性毕业生通常比男性毕业生获得的职位**更差**，（3）**声望的提高会使某一机构的博士毕业生获得更多、更好的教师职位，并增强其在本学科领域的影响力**。这些结果提高了我们量化和理解学术界系统结构的能力，揭示形塑个人职业轨迹的因素，并在教师聘用和社会不平等之间定义新的联系。

结论

在抽样学科中，我们发现教师产出（输出教师的人数）是**高度倾斜**的，**25%的机构里所产出的终身轨的教师数量**，占比达到 71% 至 86%（表 S2；此后的表格范围除非另有说明，都指的是给定的三个学科的比例范围）。学术单位内的教师人数（聘用的教师人数，即单位的规模）也是倾斜的，有些单位比其他单位大两到三倍。商学院规模尤其大，通常包含几个内部部门，平均有 70 名教职员工从其他样本内的单位获得博士学位，而计算机科学和历史的平均数量分别为 21 和 29（参见补充材料）。然而，学科内部的规模差异无法解释观察到的职位输出差异。如果职位只是简单地与单位的规模成比例分布，那么它们是无法出现统计差异的。对规模-职位比例分布假设

的简单检验表明它可能被拒绝[Kolmogorov-Smirnov (KS) 检验, $P < 10^{-8}$; 图 2; B 和 C], 表明不同机构输出教师的能力存在实质性的差异。

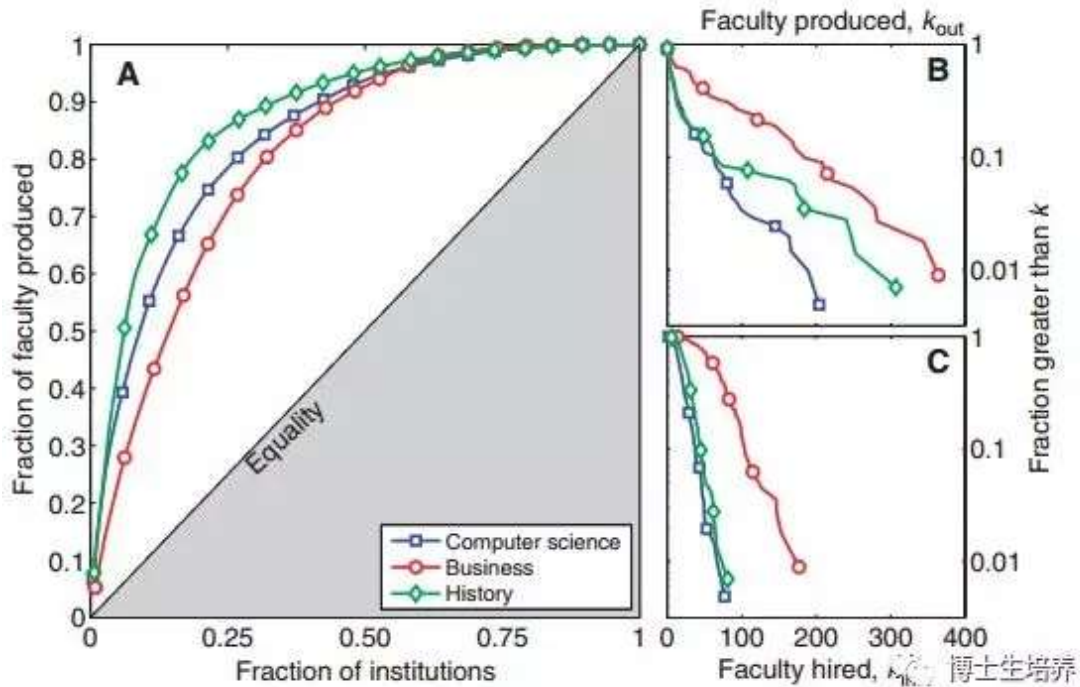


图 2 职员产出不平等

A 图是用洛伦兹曲线显示院系教师产出的不平等现象。B 图显示教师输出和引进间的互补累积分布情况。

基尼系数, 是衡量社会不平等的标准尺度, 被定义为均匀随机的观察值之间的平均相对差异。因此, $G=0$ 表示严格意义上的平等, $G=1$ 表示最大程度的不平等。我们发现教师产出的基尼系数在 0.62 到 0.76 之间 (图 2, A 和 B), 表明各学科之间存在很大的不平等[参见, 美国收入分配的 $G=0.45$ (12)]。

即使是在顶层的教师生产机构之间也存在严重的不平等: 排名前十位的单位所产出的教师数量是排名在十到二十之间的单位的 1.6 至 3.0 倍, 和排名在前二十到三十之间的单位的 2.3 至 5.6 倍。要说明以上差异反映出纯粹的才能主义 (meritocracy) 结果, 即总体学术研究的功利最优性 (13), 教师产出率的差异必须反映出学术生产力的内在差异。如果纯粹的才能主义符合事实, 那么, 博士毕业于排名前十单位的教师, 其学术产出应该是博

士毕业于排名前二十到三十之间的教师的两到六倍。以上差异程度使得纯粹的才能主义看上去不可信，这显示出非才能因素如社会地位的影响。

如果教师职位获得总体上遵循一个完美的社会等级，那么一个教师就不会被比他们所获取博士学位的单位声望更高的机构所聘用（6）。一个特定的招聘网络显示这种模式的程度可以通过识别最小违规排名来确定（14, 15），这是一个最大限度接近这个极端的等级结构。

在教师聘用网络中，每个顶点代表一个机构，每个有向边 (u, v) 代表在 v 工作的教师从 u 那里获得博士学位。然后，声望等级是顶点 π 的排名，其中 $\pi u=1$ 是排名最高的顶点。等级结构的强度由 ρ 决定，即指向下方的边缘部分，那就是 $\pi u \leq \pi v$ ，最大化所有排名（14）。同样地， ρ 是教师没有在等级中获得比他们的博士学位更好的职位的比率。当 $\rho=1/2$ 时，教师在等级结构中以相同的速率向上或向下移动，无论他们来自哪里， $\rho=1$ 都表示一个完美的社会等级结构。

起推断作用的等级结构 π 和它的强度 ρ 都是有意义的。对于大型网络，采用最大的 ρ ，通常会出现许多同样合理的排名（15）。为了提取一致的排名，在得到的 ρ 不小于当前排名的情况下，我们通过重复选择一对随机顶点并交换其排名，来对最优排名进行抽样。然后，我们通过为每个机构 u 分配一个等于其在抽样集中的平均等级的分数，将抽样排名与最大的 ρ 组合成单个声望等级结构，并且这些分数的顺序给出一致的排名（参见补充材料）。对于某些机构 u 来说，该集中的等级分布提供了对等级不确定性的自然度量。

在整个学科中，我们发现了陡峭的声望等级，其中只有 9% 到 14% 的教师的职位所在机构比他们获取博士学位的机构声望高（ $\rho = 0.86$ 到 0.91 ）。此外，提取的等级比仅仅从教师生产率中观察到的不平等程度强 19% 到 33%（Monte Carlo, $P < 10^{-5}$; 参见补充材料），表明院校机构偏好聘用来自高声望机构的博士，这一偏好是特定的。

仔细研究可发现，**这些等级通常为精英机构分配更高的级别**（但并非总是如此，参见补充材料和图 S10，其将每个学科中 60 个排名最高的机构等级可视化），并且排名越高的机构，其排名的不确定性越低（图 S3）。这些基于教师聘

用网络的排名，在评估机构声望方面也至少与权威排名一样准确：我们的声望排名与美国新闻与世界报道排名 ($r^2=0.51-0.79$, $P<10^{-17}$) 和 NRC 排名 ($r^2=0.33-0.80$, $P<10^{-11}$;见补充材料) 相关，同时这两个排名之间也相关 ($r^2=0.39$ 至 0.83 , $P<10^{-13}$)。然而，与权威排名不同，声望等级排名为不同学科的教师聘用模式提供了更多洞见。

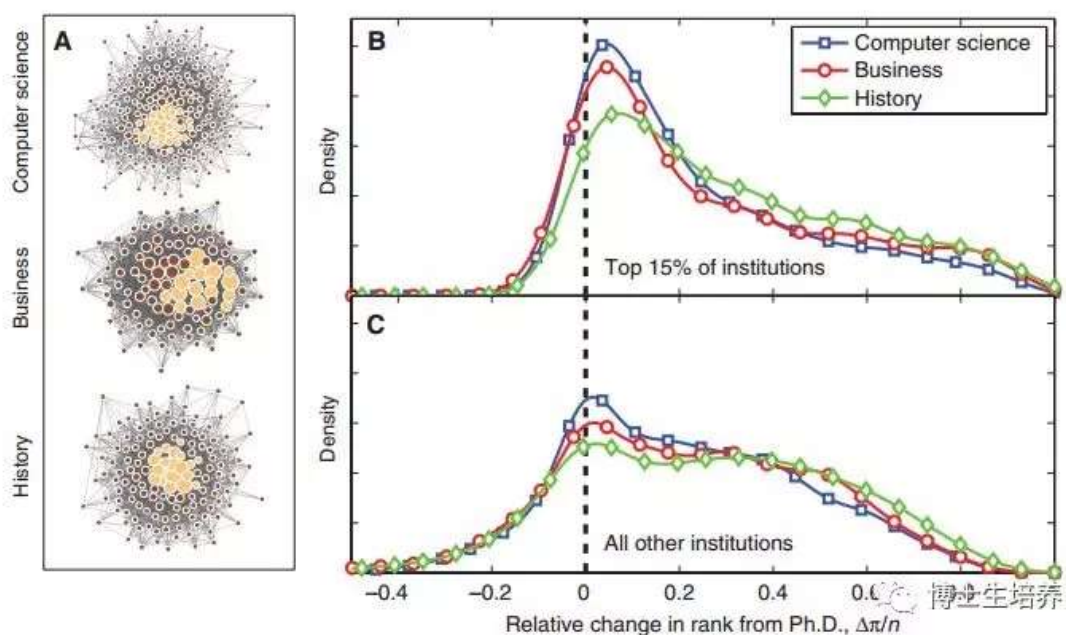


图 3 职员分配

A 图显示排名前 15% 的院系历史、商科和计算机科学教师流动网络。B 和 C 分别显示排名前 15% 的院校和全部院校的教师招聘情况可能出现的波动。

个体教师的职位经历是由相对于个人博士研究机构的排名变化所获得的。在整个学科中，我们发现教师的职位平均比他们的博士学位低 27 到 47 个等级（图 3）。中位数的变化较小，在 21 到 35 之间，表明每一个这样的分布都有相当大的右倾斜。结合观察到的不同机构中教师生产的不平等，从等级变化的平均水平来看，一个教授可以期望生产的本学科的新教师，比他们自己的博士生导师少两到三倍。教师生产的这种下降是非常陡峭的，只有前 18% 到 36% 的机构是学科内的净生产者（图 2）。

观察到的排名变化也因博士声望和性别而不同。例如，与在排名较低的院校中毕业的教师相比，在排名较高的机构中受过教育的教师，向下就业流动的比例更低（图 3，C 组；参见补充材料），表明当声望下降时，等级结

构的陡峭度增加。此外，男性和女性教师有相似但不完全相同的等级变化分布（KS 测试， $P < 10^{-3}$;图 S5 和 S6），男性的中位数变化为 21 到 35，而女性的变化为 23 到 38。对于计算机科学和商业领域中声望最高的机构的毕业生而言，性别差异最大，其中从前 15% 的单位中毕业的女性，其职位的中位数比同一机构的男性低 12% 到 18%。也就是说，这些学科中精英女性的等级比精英男性略微陡峭。相比之下，我们发现历史学科职位的中位数没有性别差异。

提取的等级制度的强度表明，单个教师职位可以仅根据博士学位机构的声望进行推断，而无需在预测模型中加入个人或机构的特征或偏好。我们通过量化和比较博士声望与替代测量的位置准确性来测试这一假设，包括权威排名和基于网络的测量。每一个度量单位都代表了一个机构的排名，从中我们计算出样本中每个教师的相对于博士等级的排名变化分布。然后，每个度量单位的预测准确性通过其在样本中的助理教授位置的曲线下面积（AUC）得分（16）来衡量。AUC 表示均匀随机真阳性（正确放置）排在均匀随机假阳性（错误放置）之上的概率。AUC 越接近 1.0，对职位的预测越准，而 AUC=0.5 的值表示准确性并不比随机高。

在各学科中，声望层级可以最准确地预测教师职位获得，AUC 变化范围在 0.58 到 0.67 之间（参见补充材料；图 S9）。其他单一的测量维度，包括美国新闻与世界报道和 NRC 的权威排名，准确性较低。此外，替代维度准确性的相对排序在各学科之间并不一致，表明可推广性不好。相反，声望始终是最好的预测因素。适度的总体准确度（AUC < 0.7）表明其他因素可能在特定的职位获得中发挥重要作用，例如，特定部门在特定年份中招聘的特定职位中出现的偶然情况。确定和量化这些因素的重要性将为教师招聘市场的效率提供新的视角。

总之，这些结果与以经典的**核心-边缘模式**组织中的学术系统大致一致（17），声望越高，越能在网络中占据更为中心的、连接更广泛的和影响力更大的位置（18）（图 4）。在支持这一结论的同时，我们发现网络中心性的标准度量与声望等级密切相关（参见补充材料;图 S8）。例如，谐波中心性（指的是从 u 到所有其他顶点（19）的平均最短路径距离的反向度量）随

着声望的提高逐渐增加，这意味着与低声望机构相比，高声望机构与其他机构之间的中介更少。因此，与处于外围的教师相比，在核心机构的教师实际上处在一个“小世界”之中（20）。

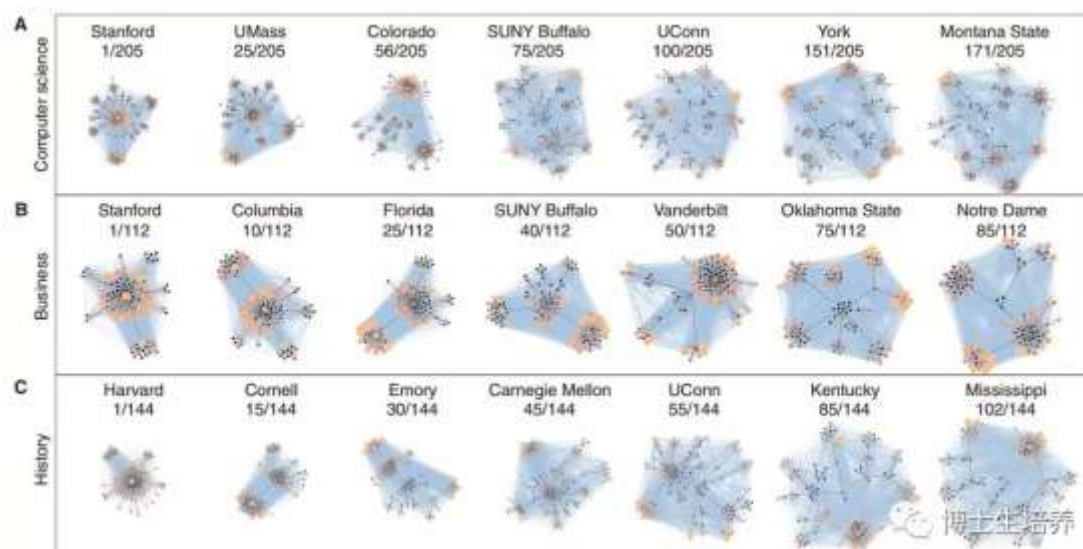


图4 核心-边缘模式

强大的核心-边缘模式对思想的自由交流具有深远的影响。研究兴趣、合作网络和学术规范通常在博士培养期间得到巩固（2）。因此，高声望机构集中化的和高度联系的位置，通过博士职位获得实现了对整个学科的研究议程、研究组织和部门规范的实质性影响（6,21）。核心与整个网络的紧密联系意味着来源于高声望核心的思想，不管其优点如何，都更容易在整个学科中传播；而来自低声望机构的思想，必须经过更多的中介过滤。以下数据更好说明高声望机构的中心性和狭隘性，我们观察到来自声望排名前 15% 的单位的教师，有 68% 到 88% 从同一排名的单位中获得博士学位，只有 4% 到 7% 的人从低于前 25% 的单位中获得博士学位。

讨论

这些结果证实了**院校声望在学术界的教师聘用中对于院校本身和寻求教职的个体都发挥着巨大作用**。通过直接依据对声望的敏感度来制定决策，或者间接改变教职人员分布情况，声誉等级也可能会影响主要研究方向、资源分配和教育产出等其他学术活动的结果。尽管在可衡量的声望范围内，实

力和社会地位的影响是混淆不清的，但所观察到的等级体系分化十分明显，以至于将其结构全部归因于实力上的差异似乎是不可信的。

计算机科学和商学中观测到的按性别统计的就业职位质量差异支持了这一结论。类似的性别不平等模式在学校评估的其他方面也被观察到了，尤其是在科学当中，它表明非才能因素发挥了系统性的作用。**相较而言，历史学科中的教师职位获得没有表现出这种性别不平等。**这一差异是否与历史学科的男性教师比例（64%）较计算机科学（85%）和商学（78%）而言较低有关，尚不得而知。找出影响这些差异的机制可能会进一步阐明性别不平等的起源以及其他非才能因素在教师聘用中的作用。

值得注意的是，尽管计算机科学，商业和历史学科实践和评估标准之间存在广泛的差异，但这些学科呈现出定性和定量相似的模式。这种共同的结构表明，强大的声望等级可能是根本性的，一些其他使用不同评估方法的针对单一学科的研究也得出了类似的结论，这些研究涵盖数学、经济学、法学、社会学、政治学和组织科学（8,25-31）。但是，产生和维持这些等级的具体机制仍不清楚（见补充材料）。在更好地理解这一机制的性质之后，我们就能在声望等级体系中将真正的才能从纯粹的社会地位中剥离出来，并为当前教师市场的运作提供新的洞见。

在我们的分析当中，**机构声望既取决于教职生产数量，也取决于职位质量。**一些机构通过成功地将数量较少的教师安排在排名靠前的机构中来获得相对较高的声望。例如，在计算机科学领域，加州理工学院排位高于 98.5% 的其他机构，但他们输出的教师数量却少于另外 27 所排名其后的院校。

这些不同寻常的成功机构，以及 9%-14% 的职位声望超过其博士机构声望的教师都提出了一个尚未解决的难题，即究竟是哪些特征将他们与更为典型的案例区别开来。如果存在一些区别这些特殊教师的因素，并且知道这些因素在何种程度上弥补了他们博士毕业院校声望不高的不足，认定这些因素会对学术界教职人员招聘机制具有重要意义。妥当研究这一问题，需要有关研究特征、学术指导、机构资源等方面的详细数据，这都是本研究所没有涉及的。

对所有学科全部教师的职位进行全面研究将是一项实质性的工作，但也能了解这些不寻常的机构和个人如此成功的原因提供广泛的基础。这样

一项广泛的研究还将有助于更全面地了解影响跨学科教师流动的过程、新领域的形成以及跨学科研究的新兴实践。

在机构层面，基于教师招聘网络的评估为美国新闻、世界报道和 NRC (10,11) 等广受批评的方法提供了原则性和数据驱动的替代方案。从教师招聘网络中提取的声望层级不是选择任意因素的任意加权，而是使用许多半独立专家组（教师本身）对研究和教育结果的集体评估。此外，教师招聘是一个高成本且高度分散化的过程，其成果通常随时可供公众使用。这些因素表明，声望等级所依赖的数据可能更具稳健性，能够避免机构对数据的人为操控，众所周知，美国新闻和世界报道的排名就存在这一问题。

更广泛地说，不同学科教师职位中存在的严重社会不平等引发了几个问题。**多少才华横溢的学者，其学术生涯轨道会因为就业市场偏好博士学位的声望等级而脱轨？**如果不同机构输出教师能力的差距变小，学术界在总体上的学术进步是否会更大？如果更低程度的不平等会推进总体的学术进步，什么样的变革在实践中是更加利大于弊的？这些是关于学术系统的结构和效能的复杂问题，需要进一步的研究来回答它们。不过我们也注意到，关于经济以及收入和财富不平等的研究可能会对严重不平等的实际后果提供一些见解。

最后，在我们的网络分析中，教师招聘没有任何特殊情况，从交互数据中提取声望等级的相同方法，可以用来研究其他形式的学术活动，例如机构之间的科学引用模式 (32)。这些方法也可以被用于分析公司内部或商业部门之间的员工流动，从而可能揭示经济和社会流动的机制 (33)。最后，因为研究生项目接纳其他院校的毕业生，所以也可以采用类似方法来评估本科生课程的教育成果。