



重庆大学

经济与工商管理学院

Economics and Business Administration

INNOVATION NETWORK

创新网络

郑维伟 郑楠 夏小宁

重庆大学经济与工商管理学院

2020年6月30日

INNOVATION NETWORK

Technological progress builds upon itself, with the expansion of invention in one domain propelling future work in linked fields. Our analysis uses 1.8 million U.S. patents and their citation properties to map the innovation network and its strength. Past innovation network structures are calculated using citation patterns across technology classes during 1975-1994. The interaction of this pre-existing network structure with patent growth in upstream technology fields has strong predictive power on future innovation after 1995. This pattern is consistent with the idea that when there is more past upstream innovation for a particular technology class to build on, then that technology class innovates more.

Source: Acemoglu, D., Akcigit, U., & Kerr, W. R. (2016). Innovation network. *Proceedings of the National Academy of Sciences of the United States of America*, 113(41), 11483-11488.

创新网络

技术进步是建立在自身之上的，随着一项领域中发明的扩展，可以推动相关领域的未来发展。我们的分析使用了180万项美国专利及其引证属性来绘制创新网络及其强度。过去的创新网络结构是使用1975-1994年间跨技术类别的引用模式来计算的。这种先前存在的网络结构与上游技术领域专利增长的相互作用，对1995年以后的未来创新具有强大的预测力。这种模式与以下观点相一致：当特定技术类别可以建立在过去的上游创新之上时，该技术类别可以进行更多创新。

Source: Acemoglu, D., Akcigit, U., & Kerr, W. R. (2016). Innovation network. *Proceedings of the National Academy of Sciences of the United States of America*, 113(41), 11483-11488.

一、背景与问题提出

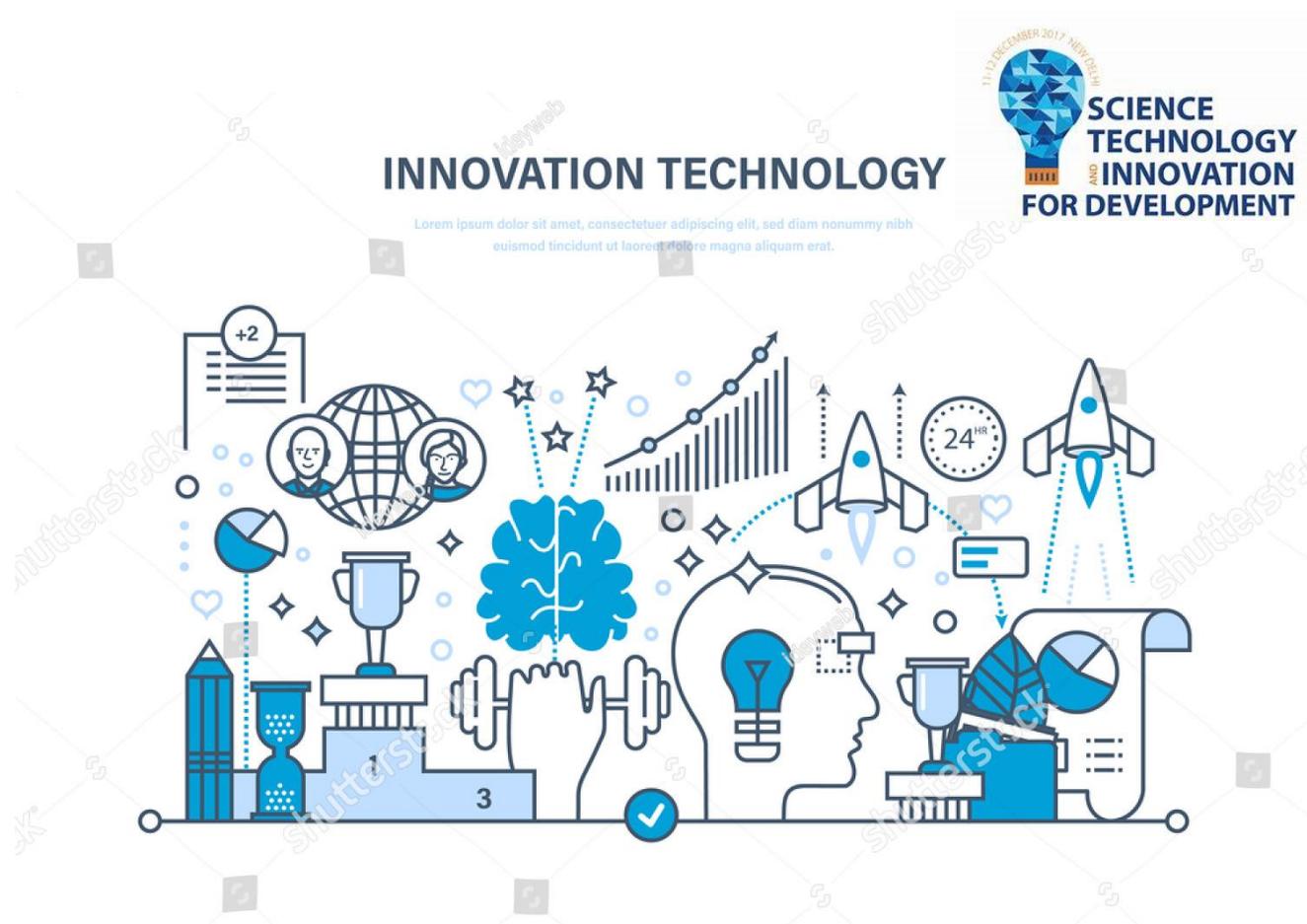
(一) Technological and Scientific Progress

技术和科学进步推动经济增长和长期福祉
(Technological and scientific progress propels economic growth and long-term well-being)。

著名的理论将这一过程描述为一个累积的过程，其中新的创新(new innovations)建立在过去的成就之上，牛顿的描述性短语“站在巨人的肩膀上” (Romer, 1990; Aghion & Howitt, 1992)。

知识发展(knowledge development)是植根于个体科学家、研究机构、私营部门行为者和政府机构的环境之中的，它们决定新发现的基本速度和方向(shape the fundamental rate and direction of new discoveries)。

现有文献对于一个技术领域的进步如何与上游技术领域的先前进步联系在一起的理解有限。



问题提出

- **跨技术领域共享知识的长期稳定性?**
- **The long-term stability of how knowledge is shared across technological fields?**
- **知识转让的速度和时机?**
- **The pace and timing of knowledge transfer?**
- **上游领域之间的密切联系如何才能对关注的技术产生实质性影响?**
- **How closely connected up-stream fields need to be to have material impact on a focal technology?**

二、研究思路与发现

用经济学中标准的内生增长和技术进步模型来推导我们的构想，假定**新思想(new ideas)的生产函数**为：

$$\Delta N(t) = f(N(t), R(t))$$

其中 $N(t)$ 是思想的存量， $\Delta N(t)$ 是产生新思想的增量， $R(t)$ 是用来产生新思想的资源(例如，科学家)。

尽管一些研究估计了思想存量 $N(t)$ 对新思想流动的影响（如是否有持续增长的回报或“挤出”外部性(“fishing out” externalities)），但大多数文献将**每个领域（它们自身的或跨所有领域的）的知识存量作为新思想生产函数的投入**。

本文通过对创新网络的研究，进一步打开黑匣子，并测量了**新思想创造对现有知识存量的异质性依赖**。

假设新的技术创新 $j \in \{1, 2, \dots, J\}$ 依赖于一个创新网络中过去所有其他领域的创新。为简化起见，资源变量 $R(t)$ 取线性形式，记

$$\Delta N_{J \times 1}(t) = \alpha \cdot M_{J \times J} \cdot N_{J \times 1}(t)$$

其中 $\Delta N_{J \times 1}(t)$ 和 $N_{J \times 1}(t)$ 分别是 t 期时 J 种创新类别所构成的 $J \times 1$ 维的创新率和知识存量； $M_{J \times J}$ 为该创新网络中某个创新类别依赖于其他类别知识存量的程度。

给定标量 α ，以及我们对技术相对增长的关注，我们可以将 $M_{J \times J}$ 的行总和标准化为 1。

$M_{J \times J}$ 中的所有项等于 $1/J$ ，代表新的创新对称地依赖于整个经济领域的技术存量；

身份矩阵(identity matrix) $N_{J \times 1}$ 则给出了仅取决于自己领域知识存量的情况。

研究思路

本文分析了**美国专利商标局(USPTO)**在**1975-2009**年间授予的**实用专利**。

每个专利记录提供有关该发明的资料(例如, **技术分类、构成当前发明的专利的引用**)和提交申请的发明人。

技术通常分为三个**层次结构**: **6个类别、36个子类别和484个类别**。

这个矩阵列出了子类别和它们的父类别, 我们的实证分析考虑了**子类别(subcategory)**和**类别级别(class-level)**的变化差异。

本文分析了**1975-2004**年间**申请的180万项专利**, 其中(需要)至少有一名发明者居住在美国城市群。

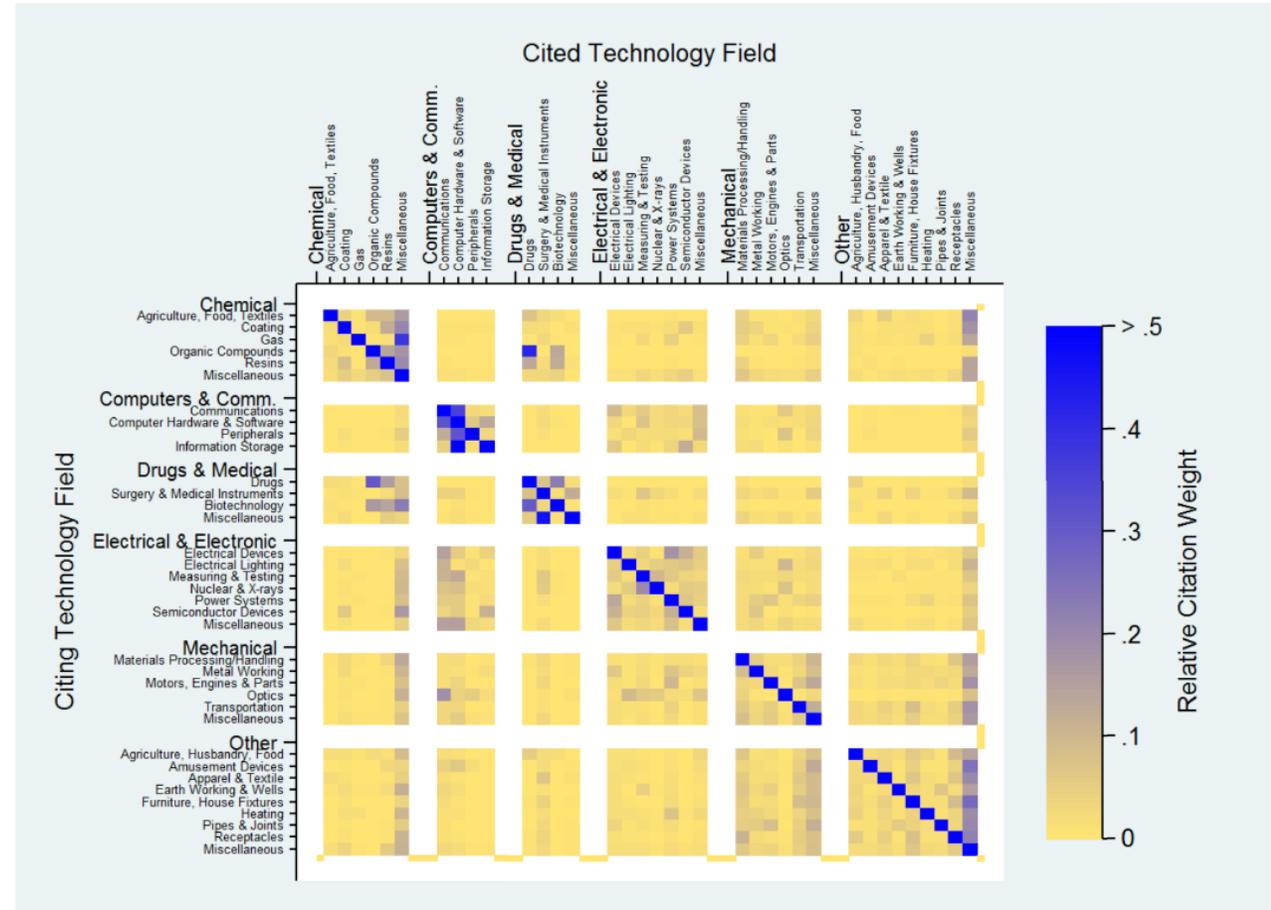
2004年的结束日期为专利审查提供了**5年的时间窗口**。在我们的数据中, **98%的专利审查**都是在这个窗口中完成的。

研究发现：矩阵形式的创新网络

本文展示了一个**稳定的创新网络**作为这种**技术和科学进步的累积过程**的管道。

分析了**180万项美国专利及其引用属性**，以**映射创新网络及其强度**。

Appendix Figure 1a: Citation matrix **1975-2004**. Extends Figure 1 to **use all citation information** over 1975-2004.



研究发现：矩阵形式的创新网络

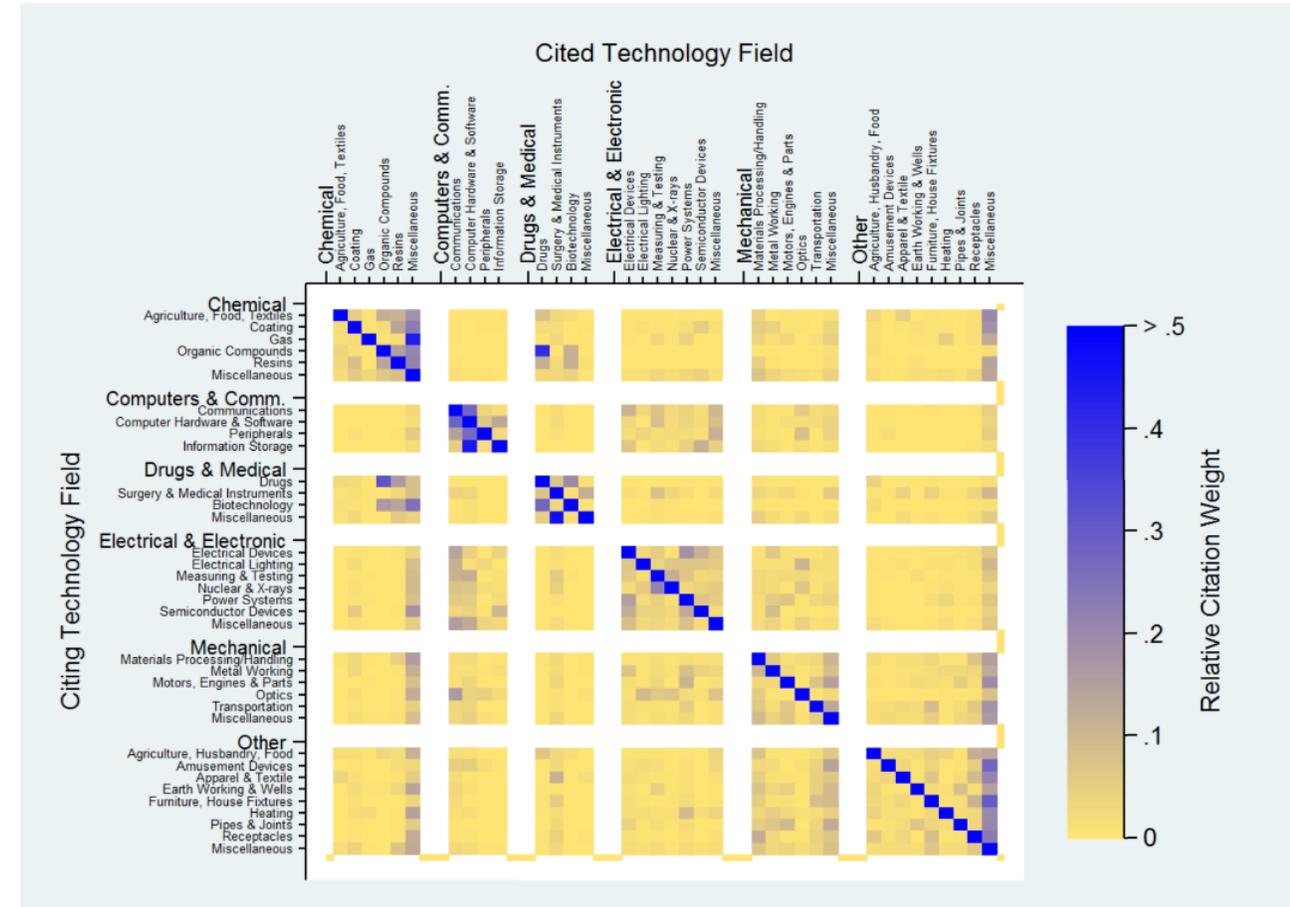
本文展示了一个**稳定的创新网络**作为这种**技术和科学进步的累积过程**的管道。

分析了**180万项美国专利**及其引用属性，以映射**创新网络及其强度**。

过去的**创新网络结构**是利用**1975-1994年**期间**跨技术类别的引用模式**计算出来的。

这种**先前存在的网络结构**与“**上游**”**技术领域专利增长**的相互作用，对**1995年以后**的“**下游**”**创新**具有强大的**预测力**。

Appendix Figure 1d: Citation matrix 1975-1994.



研究发现：矩阵形式的创新网络

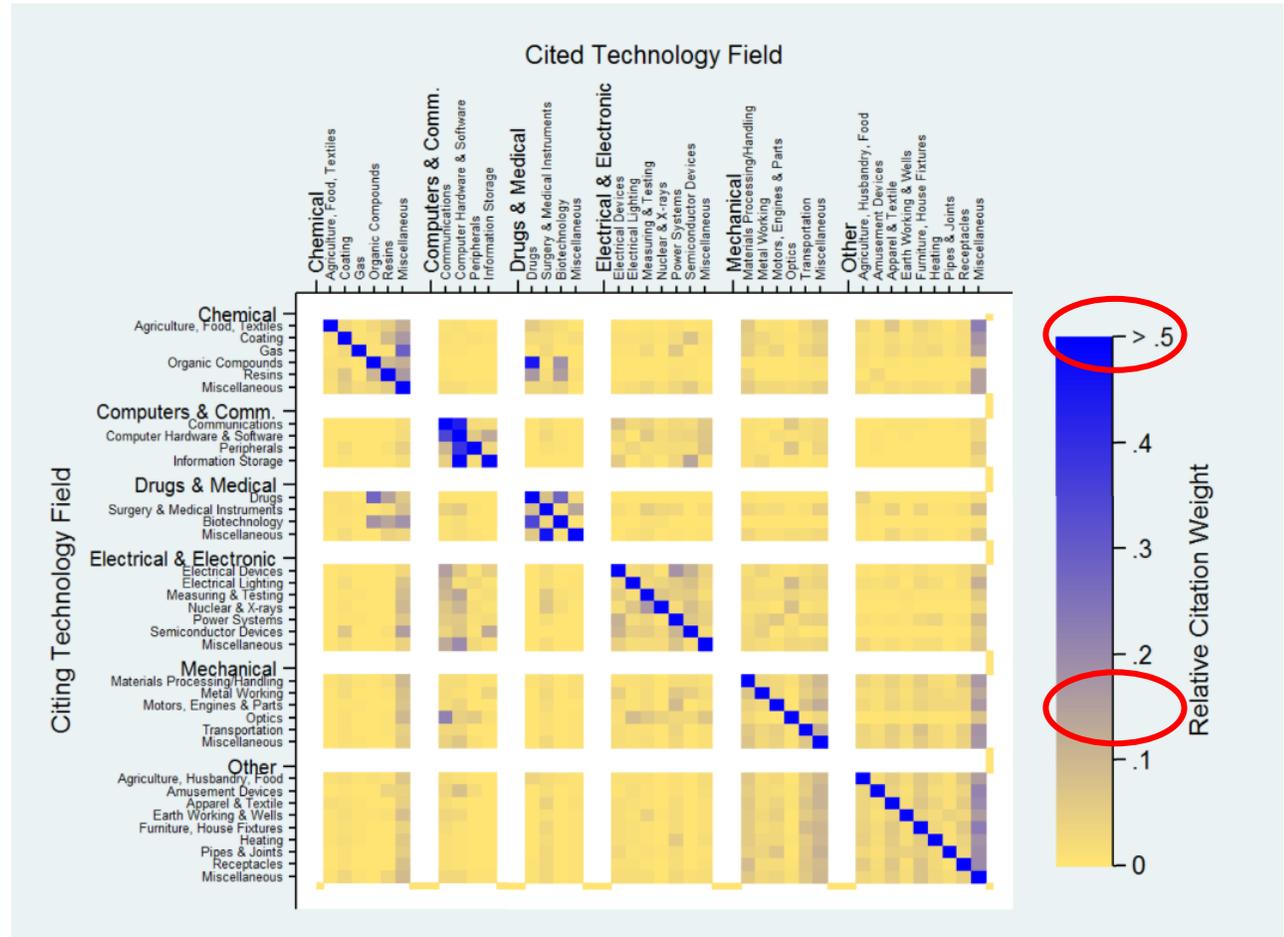
值得注意的是，1995-2004年各种技术专利水平的总体差异中，有55%可以用上游专利的差异来解释。

当使用每个领域内的面板差异时，这个解释力是14%。

在过去10年中，上游技术领域的专利增长速度更快的部门当前更有可能获得专利。

当某个特定技术类别有更多过去的创新可利用时，那么该技术类别会有更多的创新。

Appendix Figure 1c: Citation matrix 1995-2004.



研究发现：矩阵形式的创新网络

"Chemicals: Coating"和"Nuclear and x-ray"在1975-1984年期间具有了类似的专利率。

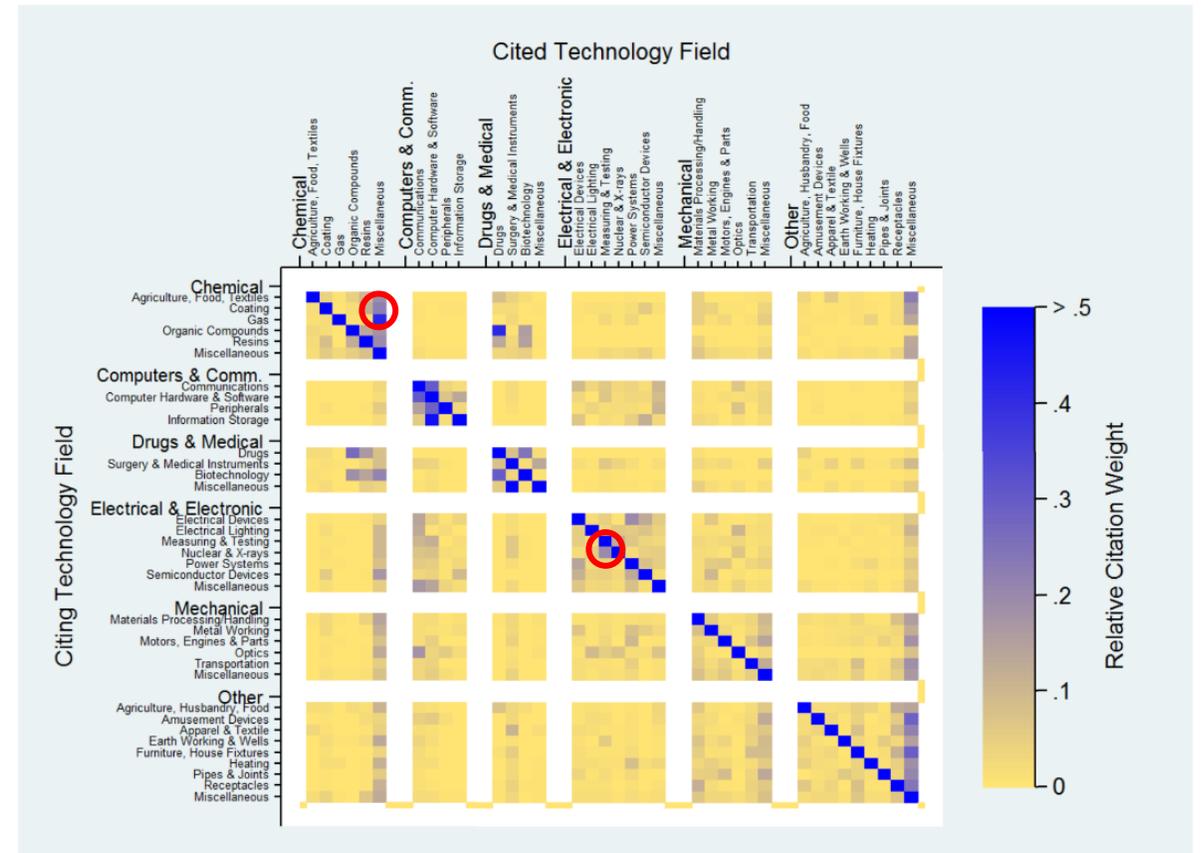
1995年之前，引用模式表明"Nuclear and x-ray"从"Electrical Measuring & Testing"吸引了大约25%的上游创新投入，而"Chemicals: Coating"对"Chemicals: Misc"也有着类似依赖。

1985-1994年期间，就新的专利而言，"Nuclear and x-ray"的上游领域专利的增长大大低于"Chemicals: Coating"。

1995年以后10年里，"Chemicals: Coating"呈现了"Nuclear and x-ray"增长的两倍。

网络异质性进一步表明，知识发展既不是全球性的，也不是地方性的，因为各个技术领域集体共享一个知识库，且每个领域只建立在自身之上。

Appendix Figure 1b: Citation matrix 1985-1994.



研究发现：矩阵形式的创新网络

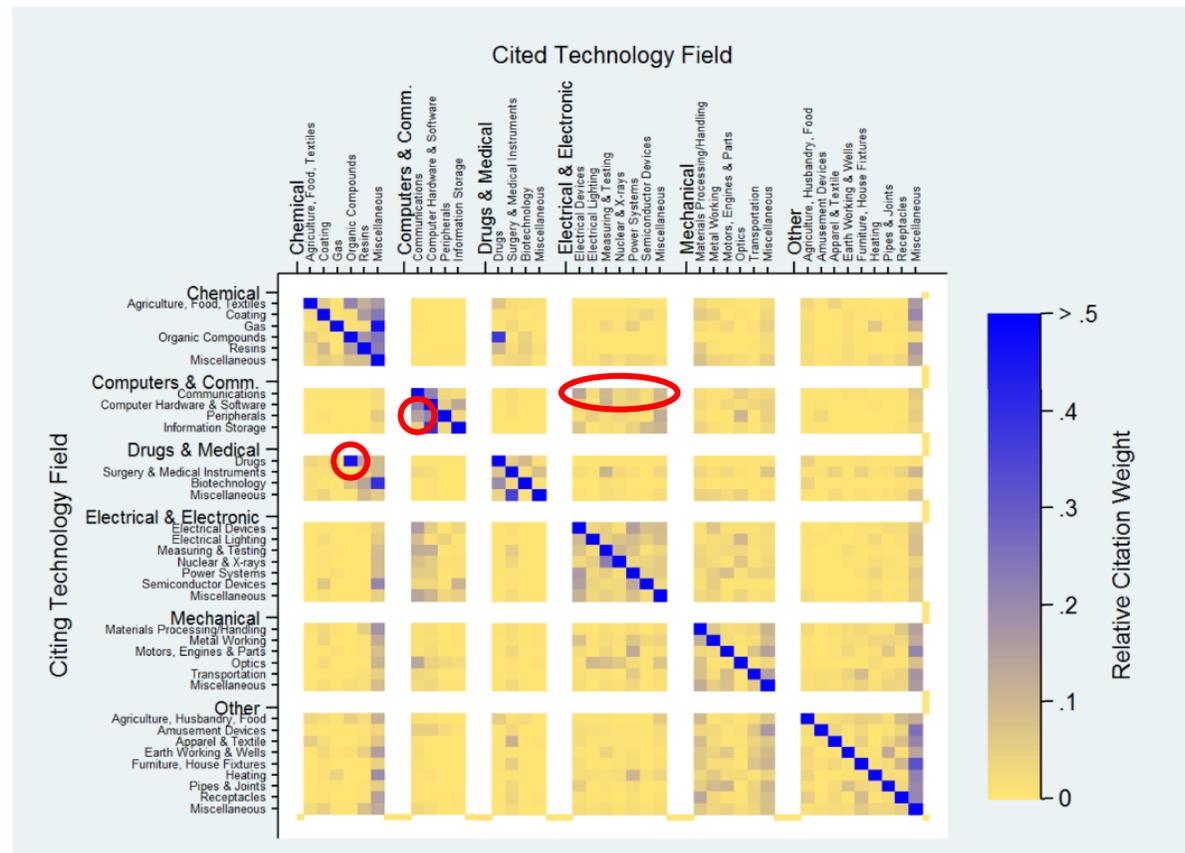
图1以矩阵形式描述了**1975-1984年的创新网络**。年份时间段是指**引用专利的日期**，前向引用专利必须在十年内。前向引用的十年时间窗口保持了每个扩散时段的观测数量的一致性。

自引(领域内引用)占大多数，并且为了视觉目的在图1中给出了一个黑色阴影。

图1显示了技术流动中的异质性。块对角线表明，**每个父类别中的子类别往往是相互关联的**，但这些流程的强度差异很大，并表现出**明显的不对称性**(如"Computers: Peripherals"中的专利往往更多地来自"Computers: Communications"，而不是相反，因为"Computers: Communications"更多地建立在电子和电子子类别的基础上)。

还有一些**跨技术类别联系的突出例子**(如"Organic Compounds"和"Drugs"之间的联系)。

Figure 1: Citation matrix 1975-1984. Each row describes the field composition of citations made by the technology subcategory indicated on the left hand side. Entries across cited technology fields for each citing technology subcategory sum to 100%. The diagonals—citations of one's own field, the majority of citations—are excluded from the calculation but given dark shading for reference. Appendix Figure 1 shows the 1975-2004 network and additional sub-periods.



困境：一个完整的增长核算包括一个领域的累积技术进步如何影响该领域自身的未来发展。事实上，自身技术的外溢效应通常是累积知识发展的最重要渠道，同时与吸收能力有关；根据吸收能力的概念，有学者就自身领域对吸收其他领域的外部知识展开研究(Cohen & Levinthal, 1990; Belderbos & Mohen, 2012)。

较难确定创新网络在单个领域内的重要性，因为随着时间的推移，一个领域的技术进步可能不仅仅与其过去和未来进步内生相关，还可能会受外部因素影响(如政府资金水平上升，动态行业条件)。

我们的网络分析通过由预先存在的网络结构调节，利用单个领域之外的上游技术进步来预测未来的创新，其贡献之一在最小化这些困难的识别挑战的实证环境(empirical setting)中证明了知识发展过程的重要性。

从两个方面考虑：

一种方法是**仅考虑外部网络，剔除自身领域内引用和内部溢出效应**，以更好地隔离网络性质。

形式上，矩阵 $M_{J \times J}$ 中的一个元素，表示从**引用技术 j (行)到已引用技术 j' (列)**是

$$m_{j \rightarrow j'} = \frac{Citations_{j \rightarrow j'}}{\sum_{k \neq j} Citations_{j \rightarrow k}}$$

在该表达式中， $j \rightarrow j'$ 表示**从技术 j 到 j' 的专利引用**，也意味着从技术 j' 到 j 的知识流动。

另外，为提供完整的增长视角，我们还报告了**包含自身领域内引用内部溢出效应的完整网络**的结果。

对于完整的网络计算，分母求和包括 $k=j$ 。

研究发现：网络形式的创新网络

图2用网络分析的方式描述了创新网络详细信息，这些群体在**二维空间中**有**更密切的联系**。
创新网络相当稳定。

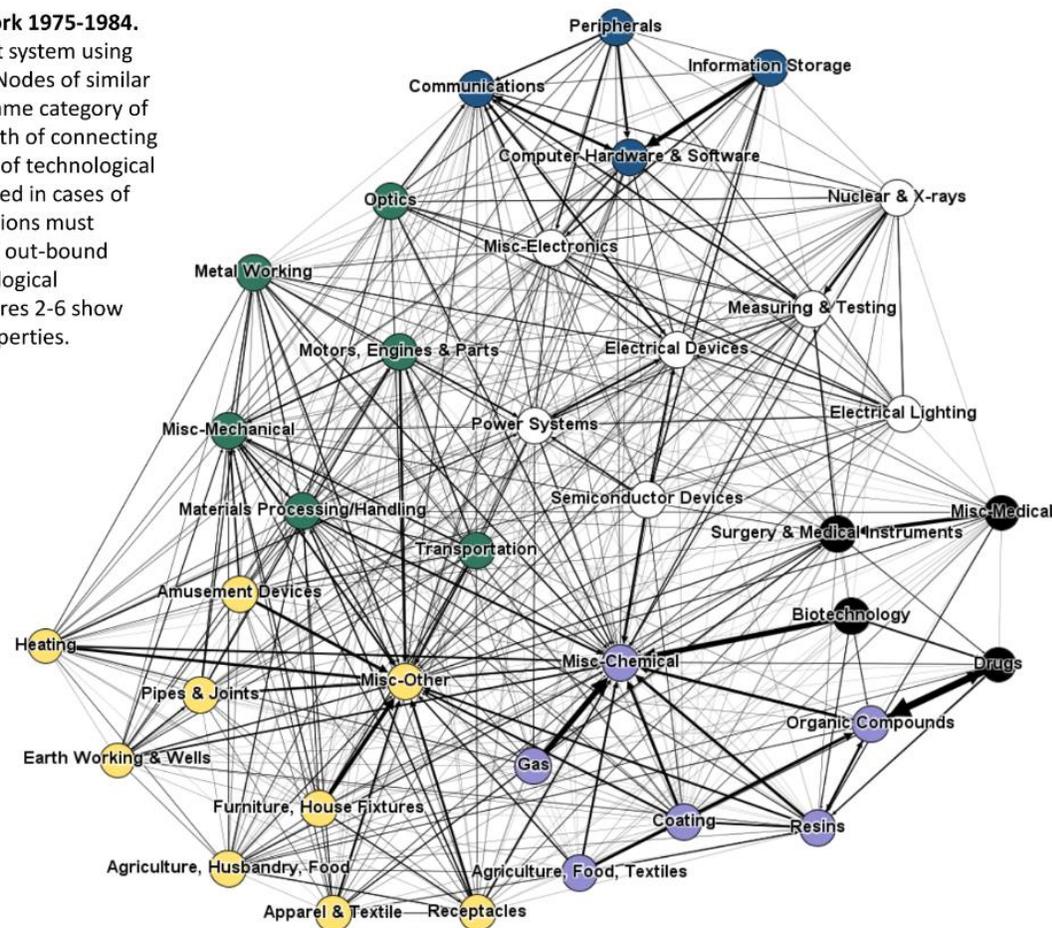
基于技术子类别的专利制度网络映射。
相似颜色的节点从USPTO系统的同一类别中提取。

连接线的宽度表明技术流动的强度，在**非常不对称**的情况下使用**箭头**。

连接必须占至少**0.5%**的外界引用的技术子范畴。

计算1975-1984年、1985-1994年和1995-2004年10年期间的 $M_{J \times J}$ ，10年期间每个“元”的相关性和秩相关性都在**0.9以上**，20年期间都在**0.8以上**。

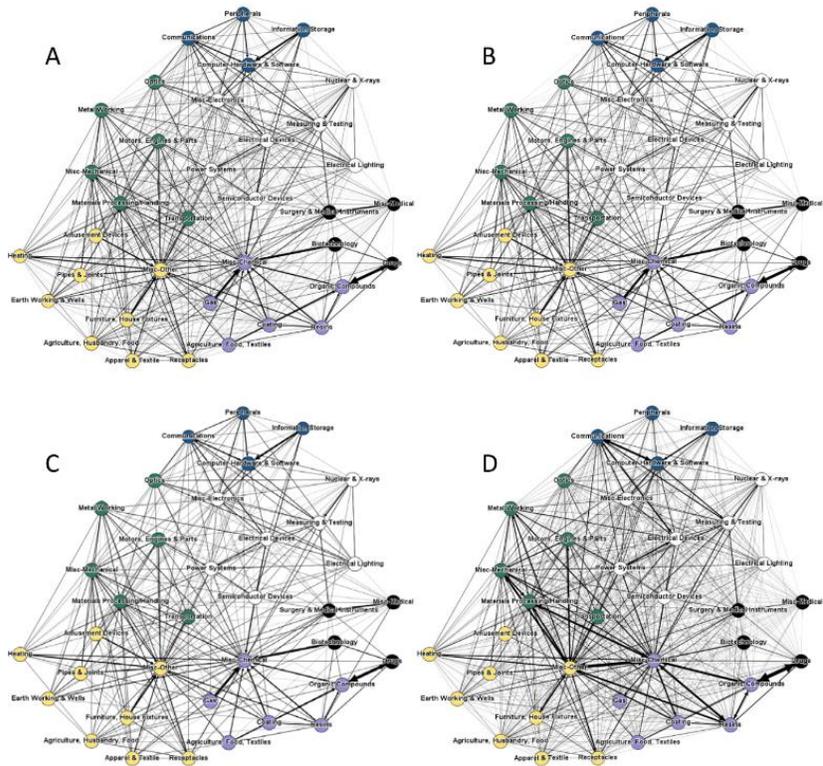
Figure 2: Innovation network 1975-1984. Network mapping of patent system using technology subcategories. Nodes of similar color are pulled from the same category of the USPTO system. The width of connecting lines indicates the strength of technological flows, with arrows being used in cases of strong asymmetry. Connections must account for at least 0.5% of out-bound citations made by a technological subcategory. Appendix Figures 2-6 show variations and network properties.



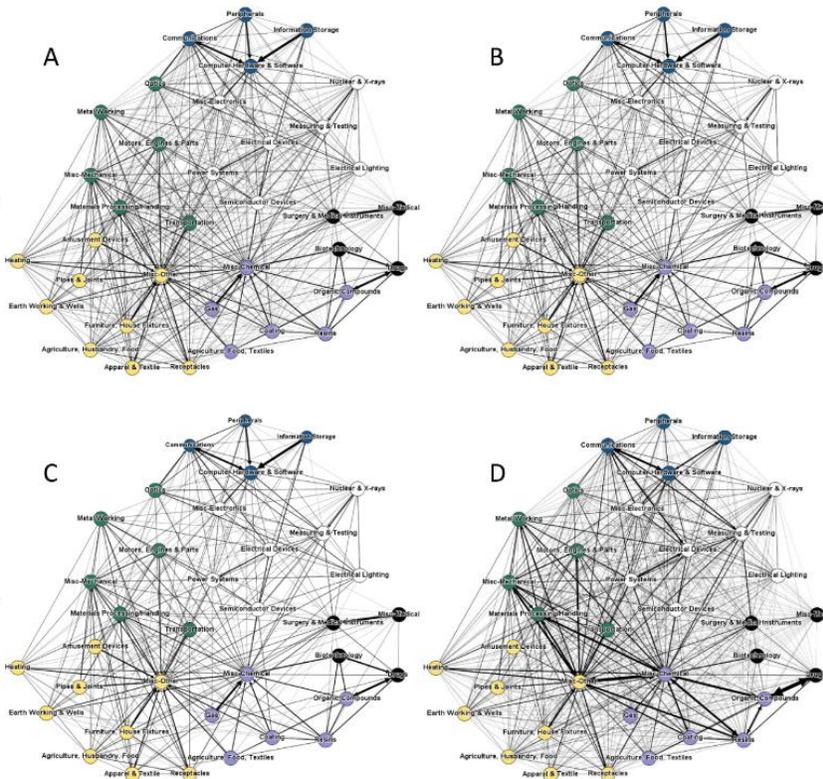
研究发现：网络形式的创新网络

附录图2-3显示了在使用更严格的阈值，包括网络边缘/连接时的可比网络结构，在未经标准化的原始引用时的可比网络结构，以及在使用更长的数据段时的可比网络结构。

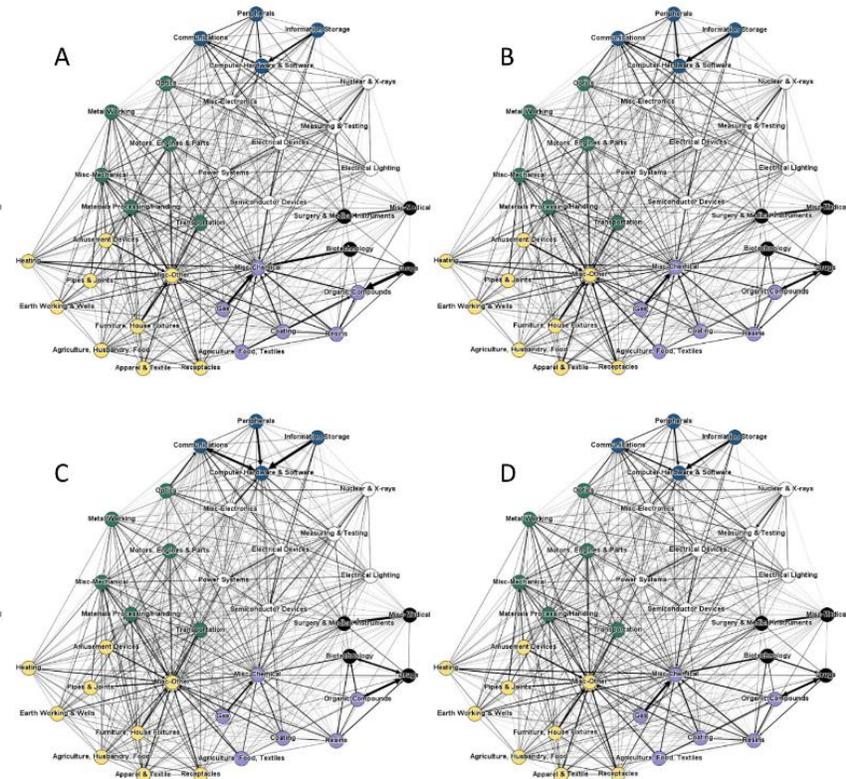
Appendix Figure 2: Innovation network 1975-1984, different edge cut-offs. Extends Figure 2 to illustrate different network cut-offs for the 1975-1984 innovation network. Nodes are colored by their primary technology category. (A) uses a 0.5% cut-off for edge inclusion, as shown in the main text. (B) uses a 1.0% cut-off for edge inclusion. (C) uses a 2.0% cut-off for edge inclusion. (D) shows a network without normalization such that every technology's outbound citations are equally weighted.



Appendix Figure 3a: Innovation network 1975-2004, full data period. Extends Figure 2 and Appendix Figure 2 to illustrate different network cut-offs for the 1975-2004 innovation network. Nodes are colored by their primary technology category. (A) uses a 0.5% cut-off for edge inclusion. (B) uses a 1.0% cut-off for edge inclusion. (C) uses a 2.0% cut-off for edge inclusion. (D) shows a network without normalization such that every technology's outbound citations are equally weighted.



Appendix Figure 3b: Innovation network 1975-2004, by time period. Extends Figure 2 and Appendix Figure 3a to illustrate different time periods and innovation network stability. (A) uses 1975-1984, as shown in the main text. (B) uses 1985-1994. (C) uses 1995-2004. (D) uses 1975-1994.

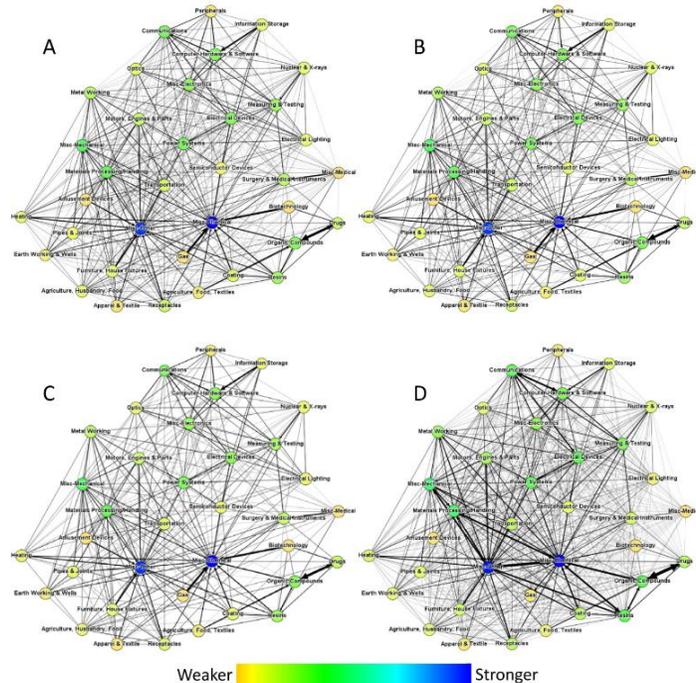


研究发现：网络形式的创新网络

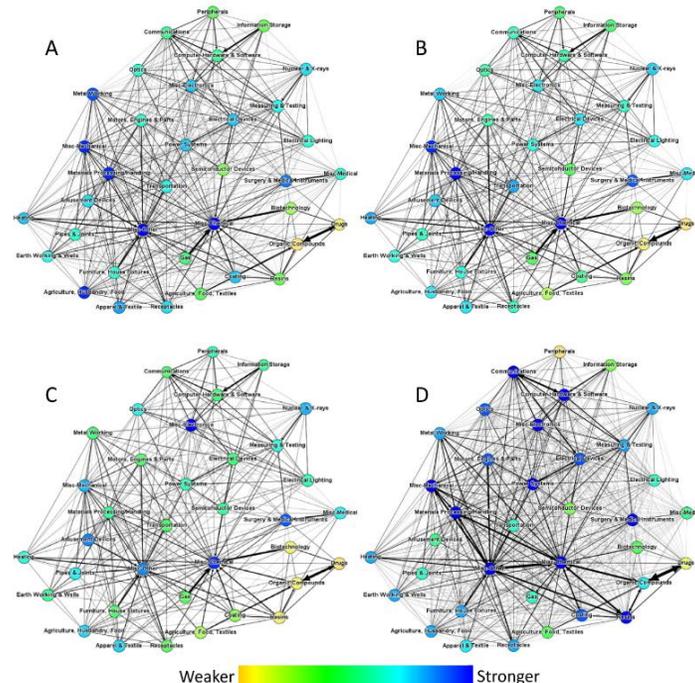
附录图4-6显示了三个经常计算的网络节点属性/指标：**点入中心度 (in-degree importance)**、**接近中心性 (closeness)** 和 **中介中心性 (betweenness)**。

一个共同的发现，许多**明显的技术领域(如"Drug")**都处于**创新网络的边缘**。**"Electrical Devices"**和**"Materials Processing/Handling"**等技术占据了**更为中心的位置**。

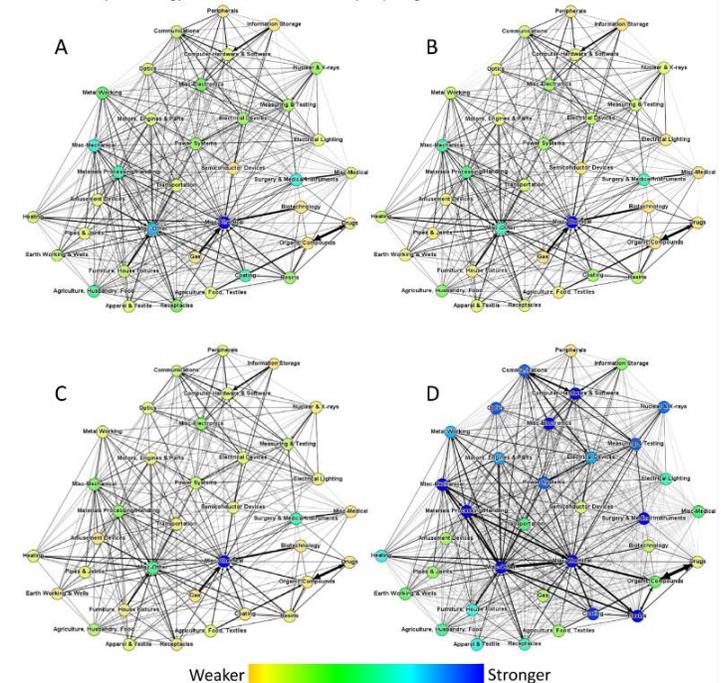
Appendix Figure 4: Innovation network 1975-1984, in-degree importance. The in-degree of a node in the innovation network is the weighted share of patents in other subcategories that are citing patents in the focal node. (The out-degree of a node is not defined in our primary case given the outbound citation normalization for each technology.) Nodes are colored by their importance using the scale at the bottom. (A) uses a 0.5% cut-off for edge inclusion. (B) uses a 1.0% cut-off for edge inclusion. (C) uses a 2.0% cut-off for edge inclusion. (D) shows a network without normalization such that every technology's outbound citations are equally weighted.



Appendix Figure 5: Innovation network 1975-1984, closeness. The closeness of a node is the average shortest-path distance between that node and all the other nodes in the network. Specifically, $Closeness_n = [1/(g-1)] \sum_m path(n,m)$, where g is the number of nodes in the network and m is any node other than n . A lower closeness score indicates a more important node in the sense of being closer on average to the rest of the nodes in the network. Nodes are colored by their importance using the scale at the bottom. (A) uses a 0.5% cut-off for edge inclusion. The closeness scores in the innovation network range from 1.06 (Misc-Chemical and Misc-Other) to 1.75 (Organic Compounds). (B) uses a 1.0% cut-off for edge inclusion. (C) uses a 2.0% cut-off for edge inclusion. (D) shows a network without normalization such that every technology's outbound citations are equally weighted.



Appendix Figure 6: Innovation network 1975-1984, betweenness. Betweenness measures how often a node links two other nodes. Identifying the shortest paths between every pair of nodes in the graph, the betweenness score of a given node n is the number of those shortest paths that include n (not including paths where n is one of the endpoints). If there are multiple different shortest paths between two nodes, the contribution to the betweenness score of a node on any of the paths is inversely proportional to the number of paths. Nodes are colored by their importance using the scale at the bottom. (A) uses a 0.5% cut-off for edge inclusion. The betweenness scores in the innovation network range from 0.21 (Information Storage) to 38.23 (Misc-Chemical). (B) uses a 1.0% cut-off for edge inclusion. (C) uses a 2.0% cut-off for edge inclusion. (D) shows a network without normalization such that every technology's outbound citations are equally weighted.



本文利用**技术专利 j' 发明后多少年内被技术专利 j 引用的知识扩散速度的异质性**，构建了我们的**创新网络矩阵**，分别对扩散过程的每一年进行建模：

$$CiteFlow_{j \rightarrow j', a} = \frac{Citations_{j \rightarrow j', a}}{Patents_{j'}}$$

其中，*Citeflow*量化了技术专利 j 在后者（ j' ）发明后的前十年中每一年引用 j' （ $Patents_{j'}$ ）的比率。

这种扩充结构扩展了用 M 矩阵描述的简单的理论模型，**允许更复杂的依赖于发明年限的知识扩散过程。**

研究思路

为了**预测未来的专利**，我们将**预先存在的网络与在关注点 t 年前10年窗口内发生的技术发展相结合**。定义 $\hat{P}_{j,t}$ 为在1994年后 **t 年时技术 j 的期望专利数**，我们对 $\hat{P}_{j,t}$ 的估计结合了前10年的专利和附加的扩散滞后 $a = [1,10]$ 年，

$$\hat{P}_{j,t} = \sum_{k \neq j} \sum_{a=1}^{10} CiteFlow_{j \rightarrow k, a} P_{k, t-a}$$

其中， $P_{k, t-a}$ 是**从 t 年起扩散滞后 a 年时技术 k 的专利数**。

例：对于**1990年申请的技术专利 j'** ，我们通过**观察前期7年扩散滞后的平均影响来模拟其对1997年技术 j 的影响**。

$\hat{P}_{j,t}$ 的**双重求和**，**遍历考虑每个潜在的上游技术类别和扩散滞后**。除了通过前期相互作用估计网络外，我们的计算要求上游专利比下游预测至少提前一年(即 $a \geq 1$)。

对于**完整的网络计算**，**第一个求和项还包括 $k = j$** 。

研究思路

使用简单方法存在一些潜在的问题。

首先，技术领域相对规模的持续存在可能导致网络重要性被夸大。

同样，所有领域的年度专利率的总体波动可能导致过分强调上游领域的重要性。

为了解决这个问题，我们考虑一个包含领域和时间控制的面板回归，

$$\ln(P_{j,t}) = \beta \ln(\hat{P}_{j,t}) + \phi_j + \eta_t + \varepsilon_{j,t}$$

其中， $P_{j,t}$ 和 $\hat{P}_{j,t}$ 是 t 年技术 j 的实际和预期专利率； ϕ_j 为专利子类别的固定效应（去掉它们的长期大小）；时间固定效应(η_t)消除了USPTO所有技术共有的拨款率的总体变化。

参数 β 的识别仅仅来自领域内的变化。

β 测度了技术 j 的实际专利申请率是否与其长期比率相关性异常过高，当它基于过去的上游创新率进行预测时。

β 显示了在控制以上条件后，**预测专利和实际专利之间的一一对应关系。**

研究发现：实际专利申请的回归情况

图3中Panel A使用子类别的年度变化报告了预测值($\hat{P}_{j,t}$)和实际值($P_{j,t}$)之间的强相关关系。

该估计结果包括通过分析1995-2004年间每年36个子类别得出的360项观察结果；每个子类别权重都是根据其最初的专利水平来确定。

考虑到外部网络，预期专利增加10%与实际专利增加8%具有相关性。我们报告的标准错误表明对于子类别中的序列相关性是稳健的。该规范解释了1995-2004年专利级别中约55%的总体变化。

完整网络估计的经验强度更强，预期专利增加10%，实际专利增加9%。

第二行估计显示， $\beta(0.85)$ 具有统计意义和经济意义。虽然估计系数小于1，但显示出预测专利与实际专利之间的密切关系。

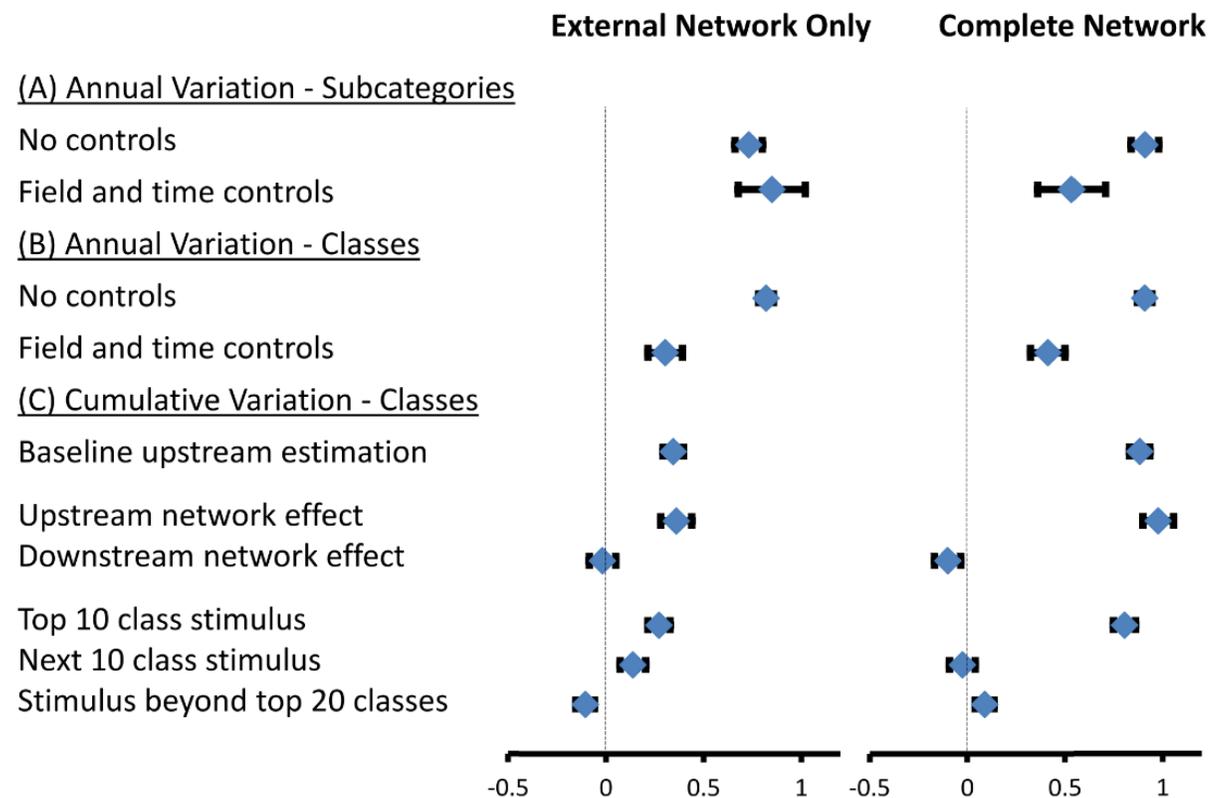


Figure 3: (A) reports regressions of actual patenting during 1995-2004 on predicted patenting calculated using the 1975-1994 innovation network and the growth in upstream technology subcategories predating the focal year. "Field and time controls" analysis reports a panel data analysis where we first remove averages from each subcategory and each year from actual and predicted values. In External Network Only analyses, we consider predicted patenting due to upstream patenting outside of the focal patent subcategory. (B) repeats this analysis for detailed patent classes maintaining over five patents per annum. (C) reports regressions using the patent class sample where we calculate cumulative actual and predictive patenting during 1995-2004 for a patent class. After reporting baseline effects in the cumulative format, we contrast the focal upstream effect with a reverse downstream effect. We next disaggregate the stimulus to demonstrate localized spillovers.

研究发现：实际专利申请的回归情况

(A)显示1995-2004年期间的实际专利申请情况，与使用1975-1994年创新网络计算的预计专利申请情况以及在关注年之前上游技术子类别的增长情况相比较。

观测值代表子类别—年配对，用父技术类别进行彩色编码，并给出45度线以供参考。

(B)面板数据分析，去除每个子类和每年实际值和预测值的均值(组内去心)，再绘制残差。

配色方案与(A)相同。

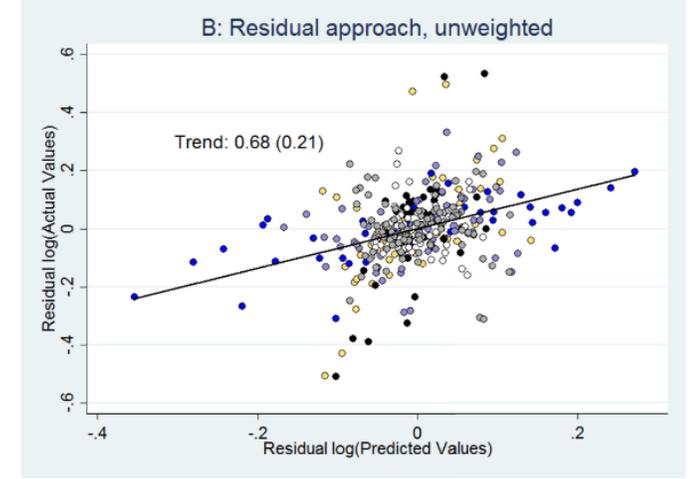
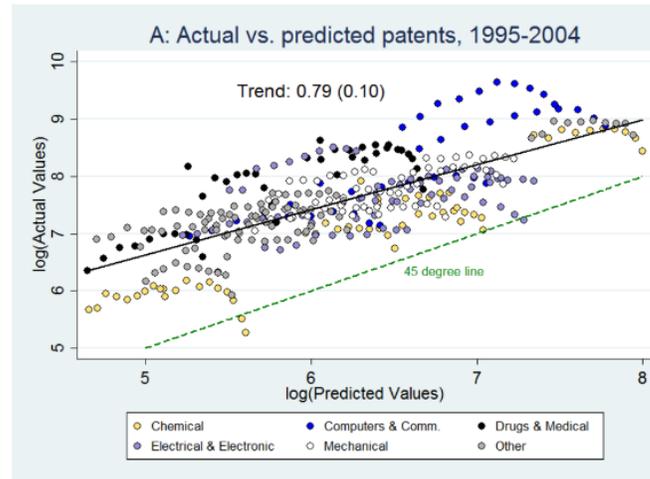
(C)展示了(B)的一个版本，根据1985-1994年期间专利申请的初始规模对观测值进行加权。

气泡大小表示权重。

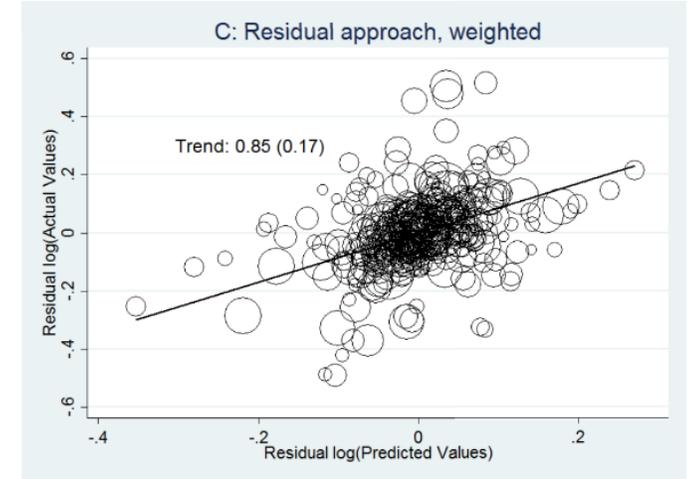
这相当于正文中报告的面板回归结果。

趋势线用子类别聚类的标准误差计算而得。

附录图7提供了这些子类别级别估计的可视化表示；结果表明我们的结果不受异常值或加权策略影响。



Appendix Figure 7: Network strength at subcategory level. (A) shows actual patenting during 1995-2004 against predicted patenting calculated using the 1975-1994 innovation network and the growth in upstream technology subcategories predating the focal year. Observations represent a subcategory-year pair, color-coded by parent technological category, and the 45-degree line is shown for reference. (B) shows a panel data analysis where we first remove averages from each subcategory and each year from actual and predicted values and plot the residuals. The color scheme is the same as (A). (C) shows a version of (B) that weights observations by their initial size in terms of patenting during 1985-1994. Bubble size indicates weight. This is equivalent to the panel regression findings reported in the text. Reported trend lines are calculated with standard errors clustered by subcategory.



研究发现：实际专利申请的回归情况

图3中Panel B使用更详细的专利类别的年度变化报告了预测值($\hat{P}_{j,t}$)和实际值($P_{j,t}$)之间的强相关关系，显示出与Panel A非常相似的格局。

在此估算中，考虑了353种专利类别，这些类别每年至少包含五项专利。

级别变化与Panel A中使用子类别发现的变化非常相似。

但Panel B的估算值较小，即预期专利申请量每增加10%，专利申请量会增加3%-4%，但在经济 and 统计上仍然十分显著。

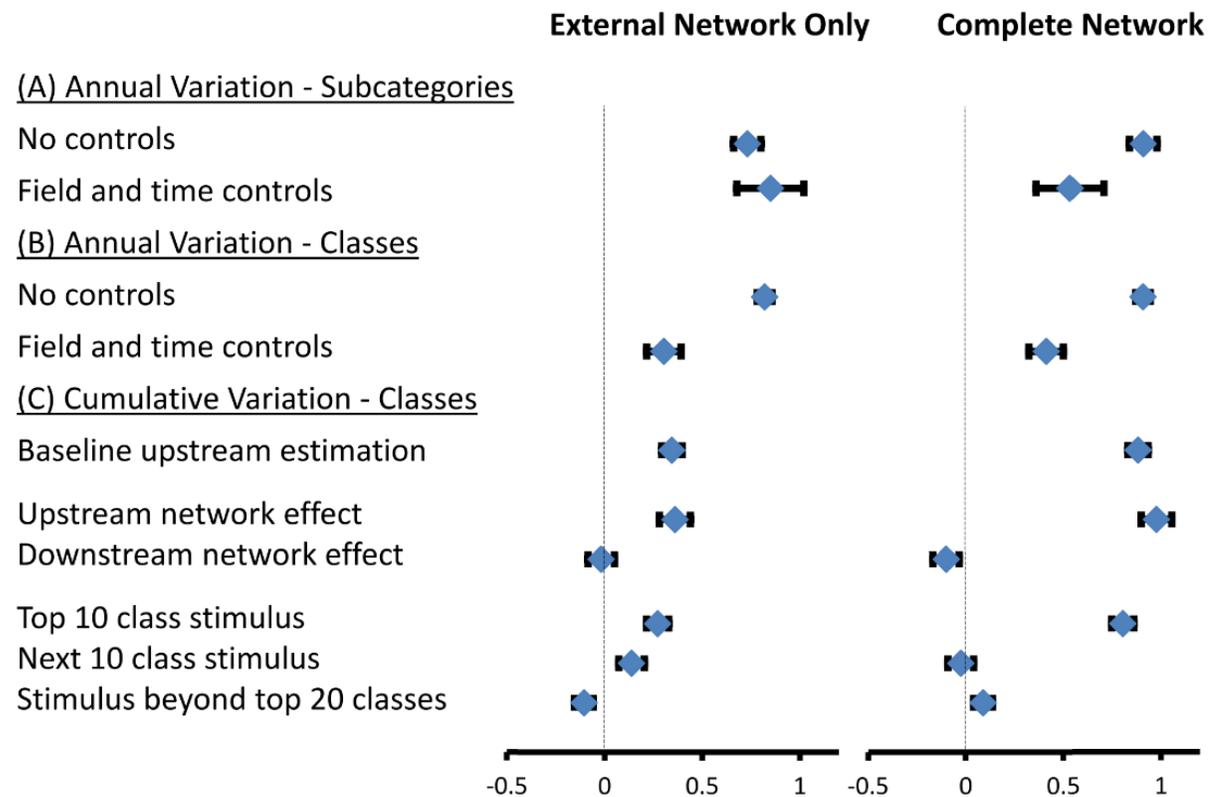
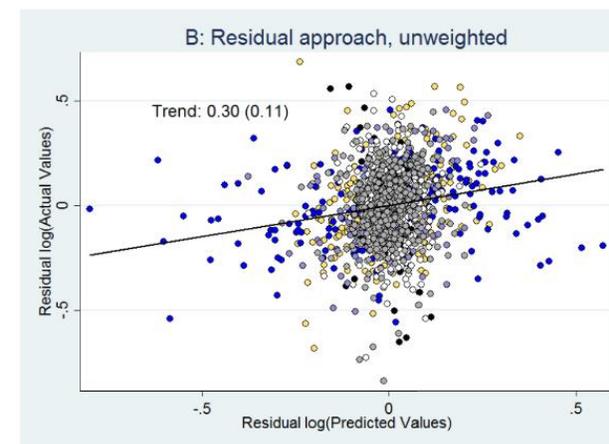
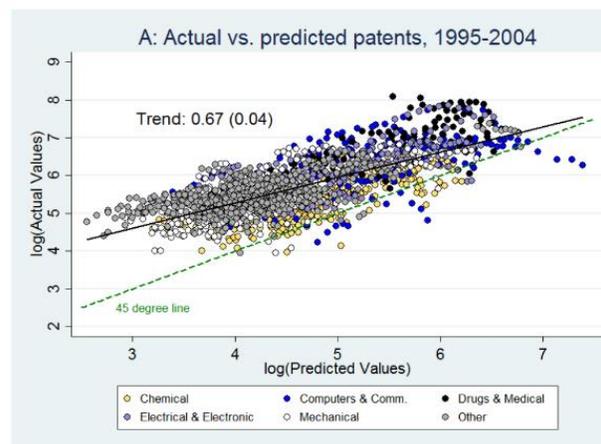
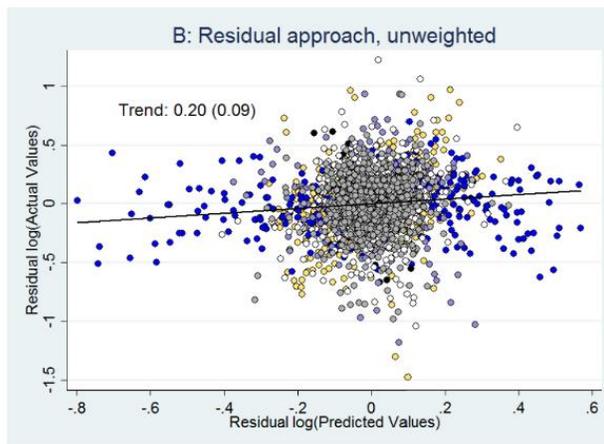
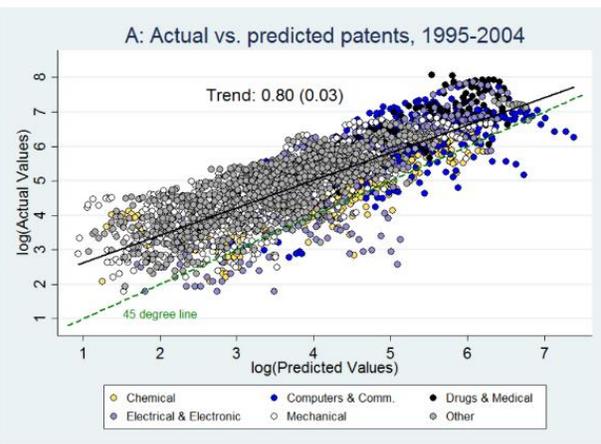


Figure 3: (A) reports regressions of actual patenting during 1995-2004 on predicted patenting calculated using the 1975-1994 innovation network and the growth in upstream technology subcategories predating the focal year. "Field and time controls" analysis reports a panel data analysis where we first remove averages from each subcategory and each year from actual and predicted values. In External Network Only analyses, we consider predicted patenting due to upstream patenting outside of the focal patent subcategory. (B) repeats this analysis for detailed patent classes maintaining over five patents per annum. (C) reports regressions using the patent class sample where we calculate cumulative actual and predictive patenting during 1995-2004 for a patent class. After reporting baseline effects in the cumulative format, we contrast the focal upstream effect with a reverse downstream effect. We next disaggregate the stimulus to demonstrate localized spillovers.

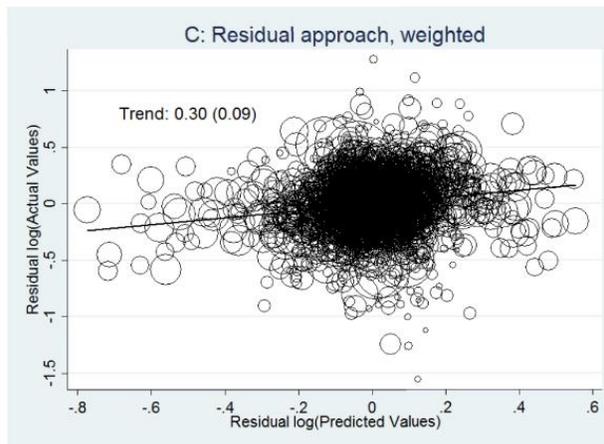
研究发现：实际专利申请的回归情况

附录图8 重复了附录7 中**1975-2004年间每年维持5项以上专利的类别**。

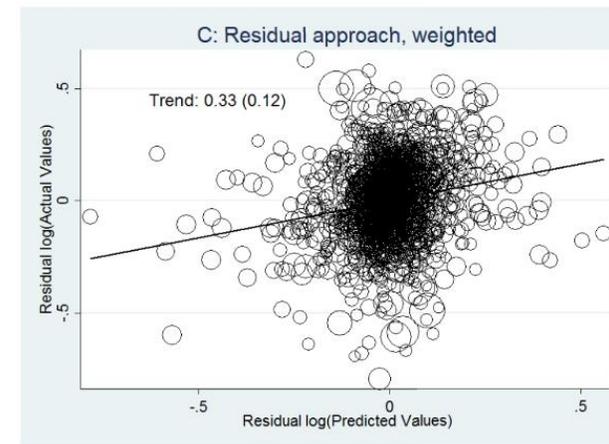
附录图9 重复了附录8中 **每年维护超过50项专利的专利类别**。



Appendix Figure 8: Network strength at patent class level. Figure repeats Appendix Figure 7 at the patent class level for patent classes maintaining over five patents per annum during 1975-2004.



Appendix Figure 9: Network strength at patent class level among large patent classes. Figure repeats Appendix Figure 8 for patent classes maintaining over 50 patents per annum.



研究发现：实际专利申请的回归情况

图3的Panel C显示了**量化创新网络强度的第二种方法**。

本文基于创新网络和历史专利级别控制，对**1995-2004年期间每个子类别的预期专利数与累积实际专利数作回归**：

$$\ln(P_j^{95-04}) = \beta \ln(\hat{P}_j^{95-04}) + \gamma \ln(P_j^{85-94}) + \varepsilon_j$$

这种方法**允许在创新网络的滞后结构影响当前的技术变化中存在更大的偏差**。

估计得到，**上游创新增加10%，对应于未来专利增加3.5%**。

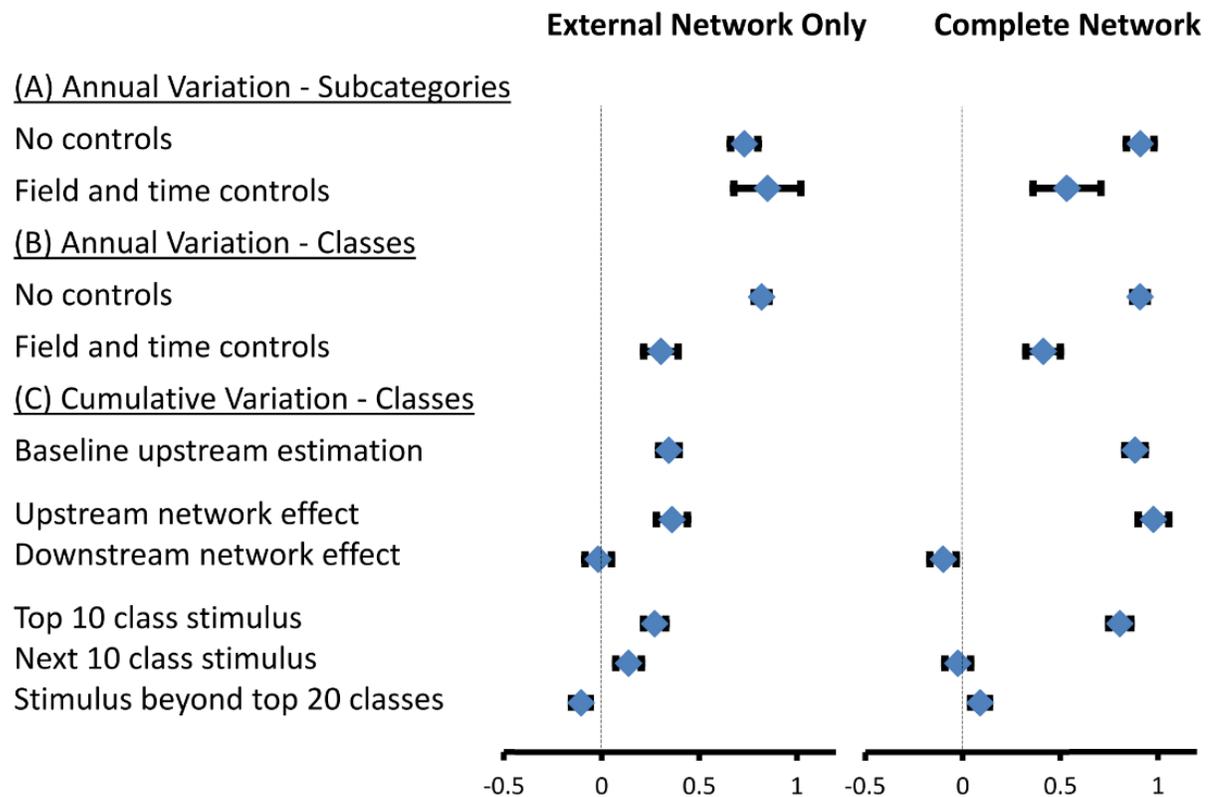


Figure 3: (A) reports regressions of actual patenting during 1995-2004 on predicted patenting calculated using the 1975-1994 innovation network and the growth in upstream technology subcategories predating the focal year. “Field and time controls” analysis reports a panel data analysis where we first remove averages from each subcategory and each year from actual and predicted values. In External Network Only analyses, we consider predicted patenting due to upstream patenting outside of the focal patent subcategory. (B) repeats this analysis for detailed patent classes maintaining over five patents per annum. (C) reports regressions using the patent class sample where we calculate cumulative actual and predictive patenting during 1995-2004 for a patent class. After reporting baseline effects in the cumulative format, we contrast the focal upstream effect with a reverse downstream effect. We next disaggregate the stimulus to demonstrate localized spillovers.

研究发现：实际专利申请的回归情况

附录图10 累积数值下的网络强度。

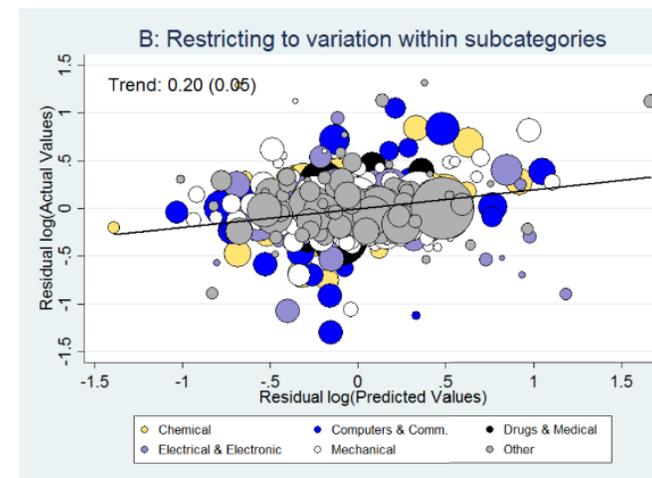
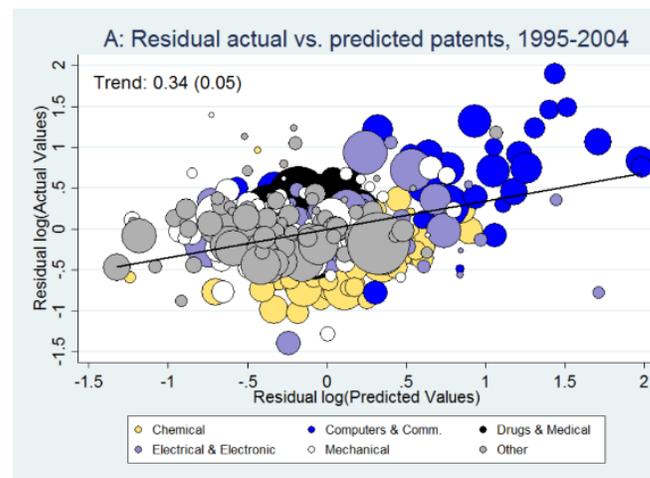
(A)我们计算了**1995-2004年按专利类别的累计实际和预测专利结果。**

图中显示了在**1985-1994年期间按类别控制初始专利水平**的回归的残差值。

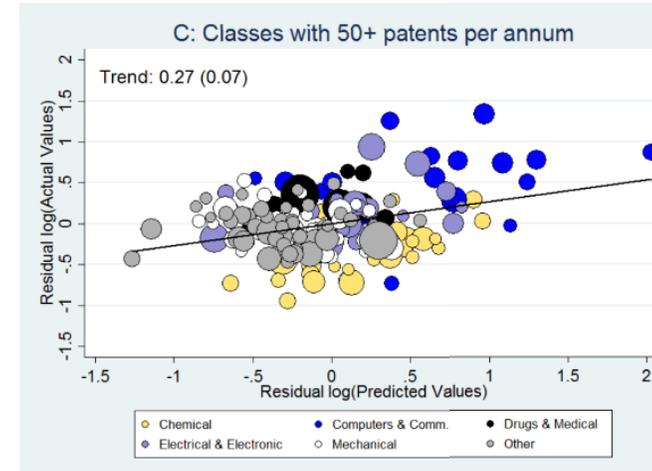
(B)显示了一个版本，在回归中控制了子类别固定效应后的残差，从而有效分离了每个子类别中的类别级别差异。

(C)将Panel (A)限定为在**1975-2004年期间每年授予50项以上专利的专利类别。**

累积方法是进行**稳健性检验**和扩展的良好思路。



Appendix Figure 10: Network strength in cumulative format. (A) shows the panel result in a format where we calculated cumulative actual and predictive patenting during 1995-2004 for a patent class. Figure presents residual values from regressions that control for initial patenting levels by classes during 1985-1994. (B) shows a version where residuals are prepared after also including subcategory fixed effects in the regressions, effectively isolating class-level variation within each subcategory. (C) restricts panel (A) to patent classes that have 50+ patents granted per year during 1975-2004.



四、稳健性检验

稳健性检验

附录表1记录了许多其他的**稳健性检验**：**控制父类别技术趋势，调整样本权重，使用增长公式，考虑二次扩散等。**

尽管它们至少取决于保留一些计算机和通信领域（computer and communication），但对于**删除任何单个子类别，其结果都非常可靠。**

在使用**国际专利分类系统**时也**同样发现这些结果。**

Appendix Table 1: Regression analysis of innovation network

	Base regression	Clustering standard errors by patent subcategory	Including patent category fixed effects	Including patent subcategory fixed effects	Using second application of innovation matrix	Using second-generation citation matrix	Using log patent class weights	Dropping patent class weights	Using constrained growth formulation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable is log cumulative patent counts in patent class 1995-2004									
A. Analysis with 353 patent classes maintaining 5 or more patents per annum									
Log cumulative patents 1985-1994	0.678 (0.061)	0.678 (0.102)	0.779 (0.047)	0.768 (0.048)	0.800 (0.062)	0.785 (0.063)	0.722 (0.057)	0.744 (0.058)	[1.0]
Log expected patenting from network stimulus	0.345 (0.052)	0.345 (0.094)	0.190 (0.045)	0.197 (0.048)	0.217 (0.054)	0.229 (0.051)	0.271 (0.053)	0.242 (0.055)	0.104 (0.028)
B. Analysis with 175 patent classes maintaining 50 or more patents per annum									
Log cumulative patents 1985-1994	0.785 (0.083)	0.785 (0.097)	0.829 (0.060)	0.739 (0.070)	0.923 (0.083)	0.915 (0.082)	0.816 (0.072)	0.827 (0.071)	[1.0]
Log expected patenting from network stimulus	0.265 (0.066)	0.265 (0.083)	0.145 (0.052)	0.184 (0.060)	0.130 (0.067)	0.133 (0.060)	0.232 (0.060)	0.223 (0.060)	0.121 (0.038)

Notes: Table quantifies the relationship between log patenting 1995-2004 by patent class and the stimulus provided by the innovation network. Estimations control for log initial patenting in the patent class during 1985-1994. Estimations report robust standard errors and weight classes by the patent count in the class during 1985-1994. Column headers report specification variants. In Column 5, second-generation stimulus is estimated through a second application of the core innovation network matrix to the predicted patenting. In Column 6, second-generation stimulus is estimated through direct measure of a second-generation citation matrix restricted to ten years.

稳健性检验

图3的底部和附录表2都量化了这样一个事实，即：**技术建立在一些关键类别上，这些类别为他们提供了创新激励因素。**

我们发现**创新与十个最重要的上游专利类别之间存在稳定的关联，其后逐渐减少。**

使用**子类别-类别结构**也显示了这种关系，尽管鉴于知识跨越技术边界，这种方法更为粗糙。

这种**网络异质性**表明，**知识发展既不是全球性的，因为各个领域集体共享一个知识库，也不是局部性的，因为每个领域只建立在自身之上。**

Appendix Table 2: Disaggregated analysis of innovation network

	Full sample with >5 per annum			Restricted sample with >50 per annum		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable is log cumulative patent counts in patent class 1995-2004						
Log cumulative patents 1985-1994	0.678 (0.061)	0.732 (0.055)	0.840 (0.066)	0.785 (0.083)	0.806 (0.069)	0.905 (0.079)
Log expected patenting from network stimulus	0.345 (0.052)			0.265 (0.066)		
Log expected patenting from network stimulus due to top 10 upstream classes		0.298 (0.056)			0.294 (0.066)	
Log expected patenting from network stimulus due to next 10 upstream classes		0.067 (0.074)			-0.041 (0.091)	
Log expected patenting from network stimulus outside of top 20 upstream classes		-0.061 (0.054)			-0.032 (0.068)	
Log expected patenting from network stimulus within subcategory			0.077 (0.029)			0.043 (0.037)
Log expected patenting from network stimulus within rest of category			0.105 (0.037)			0.097 (0.047)
Log expected patenting from network stimulus outside of category			-0.004 (0.048)			0.012 (0.054)

Notes: See Appendix Table 1. In Columns 2 and 5, we separate the upstream stimulus provided by the ten most-important upstream classes, the next ten upstream classes, and those beyond. These upstream classes are defined by citation shares made by the focal class during the pre-1995 period. In two cases, the patent class cites fewer than ten upstream categories and is excluded. Columns 3 and 6 alternatively rely on the USPTO classification system of subcategories and categories, which is naturally cruder since it is less-tailored to an individual technology's citation patterns. The disaggregated results do not add up to the total network effect due to the log transformations.

四、总结与回顾

总结与回顾

总之，本文研究发现，**上游技术发展在未来专利申请的速度和方向上扮演着重要的角色。**

更好地说明**创新网络及其不对称流动**，将有助于以更清晰的方式**对科学发现**(scientific discovery)**的累积过程进行建模。**

更好地了解这些作用可以**帮助政策决策者。**

例如，上游研究对增长非常重要的发现表明，**如果研发在一段时间内放缓，那么几年后就会感受到这种影响。**

本文在**考虑所有专利和发明**的情况下探讨了这些问题，这些专利和发明的发展可被视为**正常或常规的科学和创新。**

未来研究的一条有趣的途径是**考虑大的飞跃**是否表现出与此处所示相似的格局。

我们还认为，可以推动考虑**区域和公司层面的差异**，这可以进一步帮助我们**了解专利对经济和商业成果的因果关系。**



重庆大学

经济与工商管理学院

Economics and Business Administration

THANKS

恳请各位老师同学批评指正!