

## 一、革命与改革

### 1. 革命内涵的演变

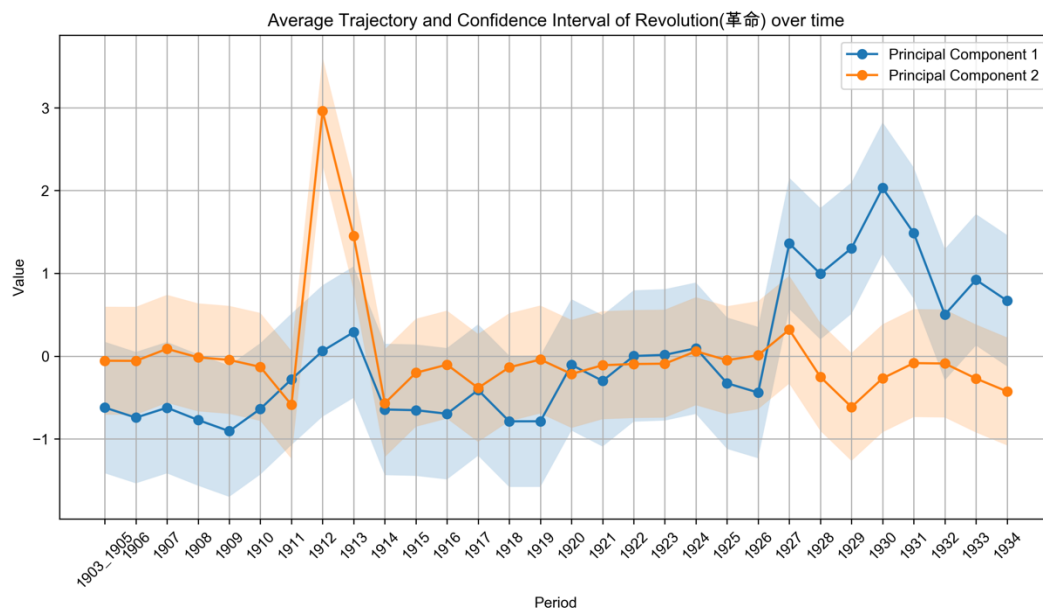
#### 1.1. 革命的语义变迁

在经过对齐变换的词嵌入空间中，我们能够通过追踪某些特定词和他们的最近邻在空间位置上的变化来理解其词义上的演变。这种演变不仅反映了单词本身含义的变化，而且还揭示了它与其他词汇关系的演变。为了更直观地展示这一变化，我们将词向量降维，并在二维空间中更为直观地展示词嵌入结果。

具体来说，我们首先将不同时间段的词向量进行对齐，通过对向量空间的旋转对齐计算，保证不同时间点的词向量能够最大程度地匹配到基准向量空间。通过动态词向量模型的对齐运算，我们保证相同的词能在空间中保持相对位置的一致性。

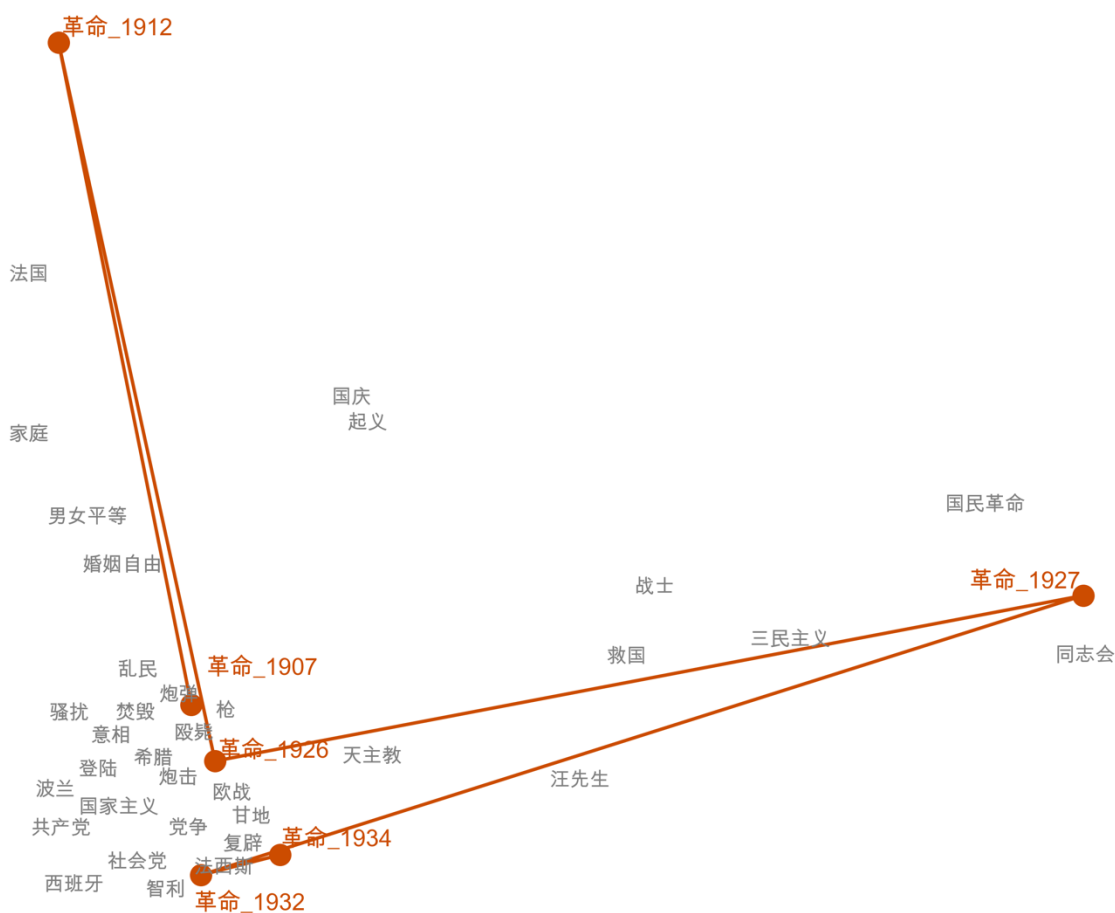
完成对齐后，我们使用 t-SNE 算法将高维的词向量降维至二维空间。我们通过概率分布的方式在低维空间中重新表示高维数据点，使得原始空间中相似的对象在新空间中以高概率保持接近，而不相似的对象则以高概率远离。

“革命”一词在语义空间上的演变十分有趣。我们提取了所有时间点上“革命”词向量的主成分，并通过降维绘制了词向量的变化轨迹，以此来展示词语随时间的总体变化轨迹。从词义变化轨迹上我们可以看到，1912 年和 1926-1927 年“革命”一词的词义出现突变，在其余年份“革命”的词义都相对稳定。



为更直观地呈现结果，我们选取了“革命”一词在语义突变年份前后的最近邻，并通过降维算法将结果在二维空间中展示。我们从最近邻的空间分布上可以看到，除辛亥革命后（1912 年）以及北伐战争（1926-1927 年）以外，“革命”一词在语义空间上聚集。这说明除这两个特殊时期以外，“革命”一词的语义都相对稳定。

Trajectories of 'Reform' through time



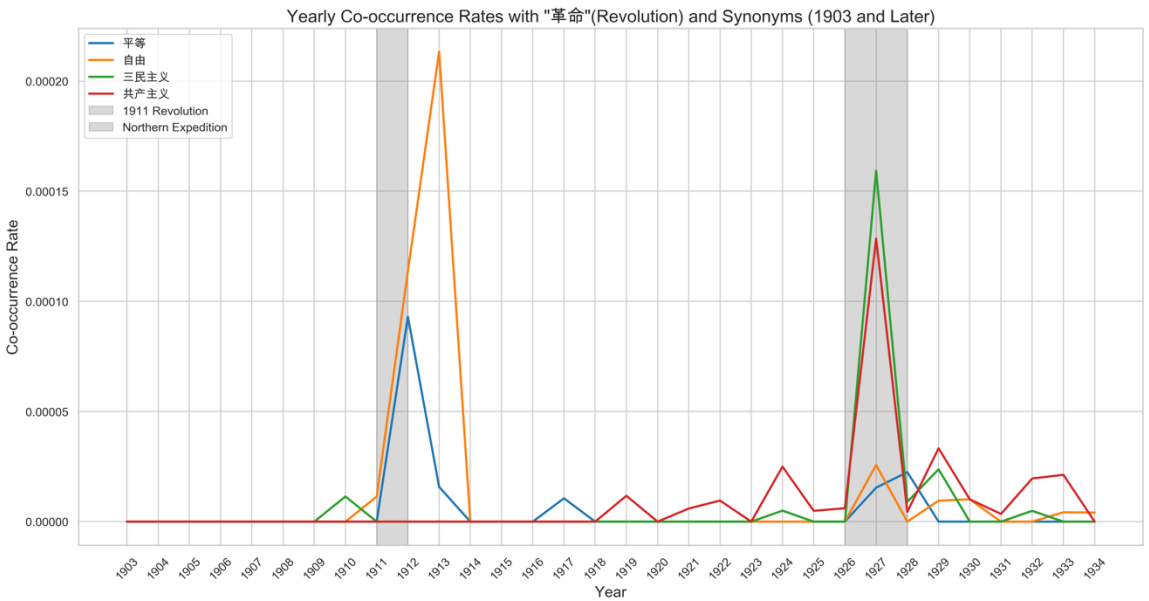
更有趣的是，在 1912 年“革命”一词在语义空间上与家庭、恋爱自由和男女平等接近，这反映了革命不仅仅局限于政治层面，也深入到社会结构和文化观念的改变。1912 年是中华民国成立的一年，这一时期的革命不仅仅是政治层面的，也涉及到社会和文化层面的深刻变革。这可能反映了当时社会对于传统价值观和家庭结构的重新审视，以及对个人自由和性别平等的追求。这也可能说明在辛亥革命期间，不同群体对于“革命”的理解和期待可能有所不同。对于一些人来说，革命不仅仅是推翻清朝，更是实现社会正义和个人自由的机会。

在北伐战争时期革命概念更加集中于政治和军事层面，这一时期国民党强调三民主义（民族主义、民权主义、民生主义）的实践，这使得“革命”一词更多地与政治意识形态和国家建设的议题相关联。另外，在北伐战争期间，国民党可能更多地通过媒体和宣传工具强调政治和军事议题，以动员和统一民众的支持。

除这两个特殊时期以外，“革命”一词在语义空间上和“骚乱”“党争”和多个西方国家有关。尤其值得注意的是，在 1930 年代，“革命”一词迅速偏离了北伐战争时期的“救国”和“三民主义”内涵，而回归到了党争的含义中。这反映了这一时期国民党与共产党关系的变化、国民党内部的派系斗争和与共产党的对立加剧，导致意识形态上的分化。这种分化可能使得“革命”一词在不同群体中承载了不同的含义和目标。而共产党的崛起和其对社会主义革命的宣传，与国民党的三

民主义和国家统一目标形成了鲜明对比。这种对立可能导致了革命概念的重新定义和解读。在大众媒体上呈现出的“革命”内涵的演变，说明政府和政治组织可能利用媒体来引导公众舆论，以支持其政治目标和意识形态。

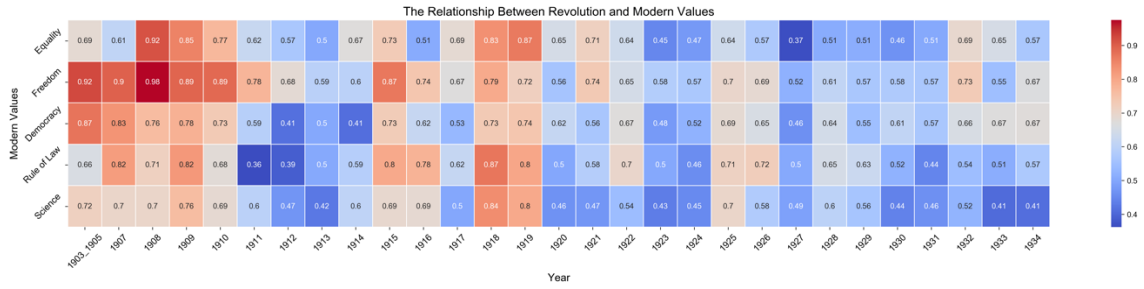
我们计算了“革命”一次与“平等”、“自由”“三民主义”和“共产主义”的共现概率。共现概率的结果与词向量空间展示的结果高度相似。在辛亥革命时期“革命”与“平等”和“自由”高度相关，这个时期的“革命”一词反映了当时社会对于民主和个人权利的渴望。到北伐战争时期，“革命”则更多地体现了政党的诉求。共现概率结果也反映了一些词向量空间未展示出的结果，1917年俄国十月革命后，共产主义思想开始在中国传播，影响了一部分知识分子和青年。1921年中国共产党成立后，共产主义与中国革命的联系更加紧密。共产主义思想与革命的联系比三民主义与革命的联系更为紧密和持久，这可能说明社会政治运动的演变，从最初的民族主义和民主主义运动，逐渐转向包括社会主义和共产主义在内的更广泛的社会变革运动。



## 1.2. “革命”的内涵与意识形态演变

### 1.2.1 “革命”对现代价值观的塑造

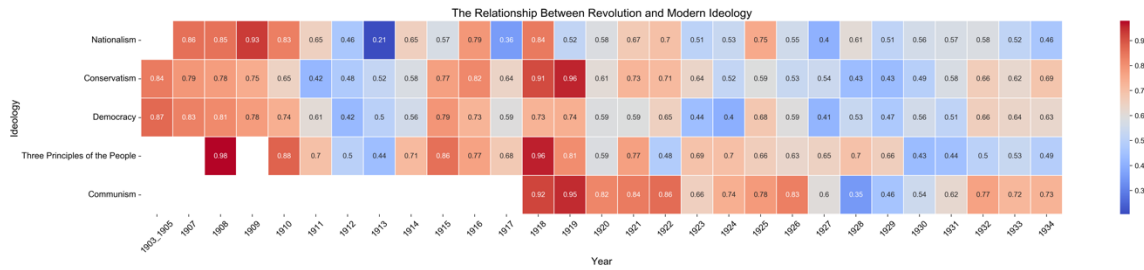
为了进一步说明“革命”在意识形态演进中的作用。我们计算了“革命”在语义空间上和衡量现代价值观的五个维度：“平等”“自由”“民主”“法治”和“科学”之间的余弦相似度。每个维度我们都选择了一组词汇，并计算“革命”与一组词汇的余弦相似度均值。得到的结果以热力图的形式呈现。从图中我们可以看出，在辛亥革命以前，“革命”一词与现代价值观的相似度极高，但在辛亥革命之后，这种相似度开始逐渐降低。在1915年到1919年，“革命”一词所代表的现代价值观内涵有所增加，但在1920年以后，这种相似度显著地降低了。



在辛亥革命之前，中国社会正经历着从传统封建社会向现代国家的转变。这一时期的革命概念可能更多地与个人自由、平等、民主等现代价值观相关联，尤其是与“自由”和“平等”的关联最强。这反映了当时知识分子和改革派人士对西方现代价值观的追求和倡导。在1918-1919年，与“革命”高度相关的现代价值观是对平等、法治和科学的倡导，可能反映了新文化运动带来的影响。

到了1920年代，现代价值观已经随着辛亥革命的成功与新文化运动的广泛开展传播开来，因此现代价值观中的革命性内涵逐渐消失。

### 1.2.2 “革命”与国家意识形态

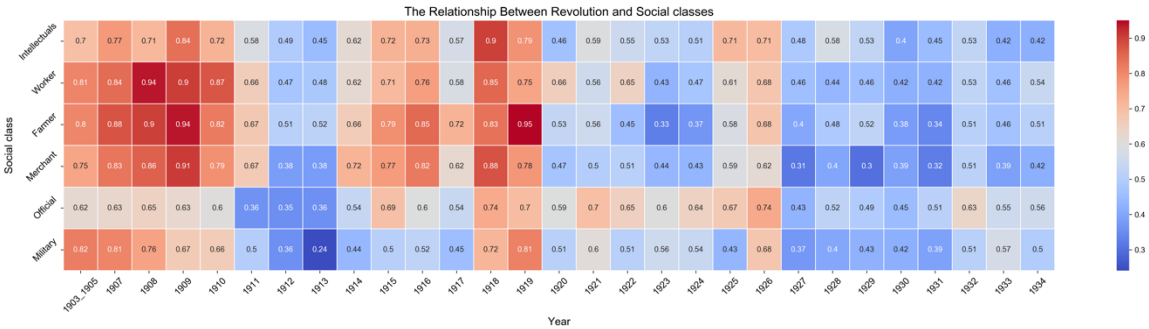


为了检验“革命”与各类意识形态之间的相关性，我们考察了“革命”与“民族主义”、“民主主义”“保守主义”和“共产主义”之间的相关性，由于这一时期的“民族、民权、民生主义”经常被概括为“三民主义”，我们同样也计算了“革命”与“三民主义”之间的相关性。从图中可以看出，在辛亥革命以及五四运动前后，社会革命与意识形态之间紧密相关。辛亥革命标志着中国从几千年的封建帝制向共和制的转变。这场革命不仅结束了清朝的统治，还催生了民族主义和民主主义的思想。在这个时期，革命不仅是政治上的变革，也是对中国传统文化与身份的一次深刻反思，与民族主义、民主主义等意识形态紧密相连。而五四运动代表了中国年轻一代的觉醒。这个时期的年轻人渴望变革，他们不仅寻求政治和社会的改变，也追求文化和思想的革新。这种对变革的强烈渴望与各种意识形态的传播相结合，共同塑造了这一时期的政治和社会运动。

另外，与之前的结果一致，在整个1920年代，随着中国共产党的建立和发展，“革命”一词同“共产主义”之间始终存在着紧密的联系。而“三民主义”中的革命性内涵在北伐战争结束之后很快的消失了。尽管在1927年以后，“革命”的语义再次回到了骚乱与战争的语义空间中，但大众媒体对共产主义的负面报道将共产主义与“革命”这种彻底的社会变革更紧密地联系在了一起。

### 1.2.3. 谁在参与革命？

在“革命”重塑国家意识形态的过程中，我们同样关心谁推动了革命，在革命中哪个阶层受到了最为显著的影响。为了回答这个问题，我们构建了社会阶层的不同维度：工人、农民、商人、知识分子、官员和军人（**这里是不是可以增加一些新的阶层？**），来考察在语义空间上“革命”与不同社会阶层之间的距离。语义空间上的距离接近代表在该时期“革命”对该社会阶层影响更大，或者特定的社会阶层在该时期主导了革命。



“革命”与社会阶层之间的关系呈现出了更为有趣的结果。首先，辛亥革命显然是一场自下而上的革命，这一时期官员阶层和“革命”的关联性相当小。但到了北伐战争之前的一段时期，革命开始由精英阶层主导，“革命”一词开始与官员阶层的关联性增强，但与其他社会阶层的关联性减弱。其次，商人阶层与革命之间的关系同样值得注意，在辛亥革命之前商人与革命的关联性强，这种关联持续到了二次革命以及五四运动时期。晚清及民国初期的商人可能支持辛亥等革命活动，因为他们反对清政府的贸易限制政策和重税。支持革命提供了一个可能的更有利的商业环境。另外，五四运动呼吁中国现代化和废除过时的做法，这与寻求现代商业实践和国际贸易的商人阶层的利益一致。但在北伐战争之后，商人阶层的利益与革命领导者或许不再一致，这种关联性也在逐渐减弱，这可能是因为随着国民党政权的建立，商业利益与政治权力的结合减少，或者是因为国民党采取了更加中心化的经济政策。最后，知识分子始终是推动变革的核心力量；工人和农民的角色随着共产主义和社会主义思想的传播而增强，在1919年这两种角色和革命的关联达到了巅峰。但随着国共合作的破裂，这种关联性在1930年代急剧降低。

1.2.4 词嵌入的回归：谁塑造了革命的内涵

在本研究中，我们采用了基于词嵌入回归（ALC）技术的方法，通过结合词嵌入模型和多元回归分析，以量化历史事件对特定术语含义的影响。首先，目标词汇及其上下文被转换成向量，利用词嵌入模型捕捉每个词汇的语义信息。然后，这些向量的平均值被计算出来，并通过应用一个转换矩阵来降低常见词汇的权重，以便更准确地反映特定上下文中的语义特征。

构建回归模型时，我们引入了一个由 ALC 向量组成的矩阵 Y，并通过回归方程

$$Y = \beta_0 + \beta_1 Post\_Revolution + \beta_3' Revolution' \times Post\_Revolution + \varepsilon$$

来分析。

通过这样的回归策略，我们分析了在革命中被多次讨论的重要概念，包括“平等”“自由”“民主”“民族”“公民”“权利”和“清”。估计范数 $\beta$ 的大小代表了Y和“革命”一词之间的相似度，而范数的显著性则代表了在辛亥革命前后，在与革命相关的语境中，Y的词义是否有显著的变化。如果我们观察到 $\beta_3$ 显著向为正，说明在辛亥革命以后，Y的词义显著与革命发生关联。

为了评估 $\beta$ 值代表的影响是否在统计上显著，我们采用排列测试来计算经验p值。通过随机打乱Y矩阵中的信息，并基于这些打乱后的数据重新计算回归多次（如999次），我们能够记录每次回归后系数的欧几里得范数，并比较这些范数与真实数据下计算出的 $\beta$ 系数的大小。计算得到的经验p值是通过比较随机数据下得到的规范化系数与真实数据下得到的规范化系数的比例来估计的。如果大量随机数据下得到的规范化系数大于真实数据下的规范化系数，得到的高p值将表明历史事件对术语含义变化的影响并不显著。通过这种方法，我们能够在统计上评估历史事件是否对特定术语的含义产生了显著影响，其中 $\beta$ 系数揭示历史事件对我们关注的目标词汇的影响程度，而通过排列测试计算出的p值帮助我们判断这种影响是否具有统计学意义。

另外，为了明确辛亥革命对我们关注的目标概念造成的影响，我们设置了对照组进行回归。对照组的回归方程如下：

$$Y = \beta_0 + \beta_1 'Revolution' + \beta_2 Post\_Control + \beta_3 'Revolution' \times Post\_Control + \varepsilon$$

第一个对照组我们选择辛亥革命前后的具体时间点进行对比分析。这样的分析框架旨在排除辛亥革命之前的事件或其他潜在混淆因素对词义变化观察结果的影响。

第二个对照组我们将袁世凯复辟定义为独立的对照组，由于袁世凯复辟发生在辛亥革命之后，为了排除辛亥革命的持续影响，我们将语料库进行了划分，并对两个语料库进行了回归。这一方法使我们能够基于时间节点明确区分并分析研究中的关键历史事件。

Table: The Estimated Shifts in Discourse following 1911 Revolution and the Control Event, across concepts.

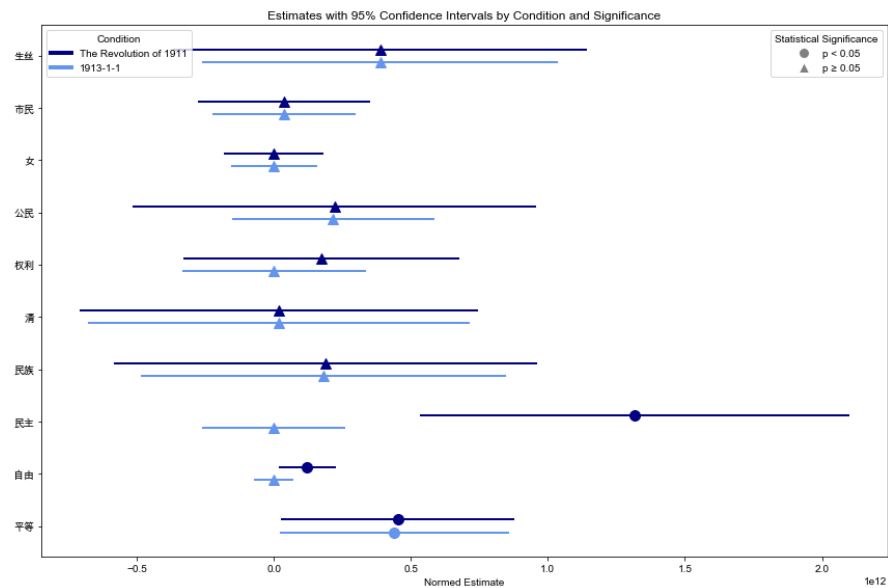
	自由	平等	民主	权利	公民
1911 Revolution	1.23*** (0.53)	4.52*** (2.17)	1.32*** (0.399)	1.73 (2.56)	2.22 (3.75)
Control1(1908/1/1)	1.21*** (0.294)	4.67* (2.67)	1.47*** (0.428)	1.74 3.25	2.85 (7.28)
Control2(1915/12/31)	0.005 (0.006)	4.21* (2.28)	0.002 (0.008)	0.002 772	2.24 (1.527)
Documents	5250	1413	495	978	3180

\*p<0.1, \*\*p<0.05, \*\*\*p<0.01



表格中的结果反映了交乘项的回归系数。我们可以看到，在辛亥革命发生之前，“自由”“平等”和“民主”等概念就开始和“革命”显著相关。但在袁世凯称帝之后，这些关键概念与“革命”之间开始减弱了。

下图更为直观地反映了这一结果，我们同样验证了在这一时期我们关注的其他概念。从结果中我们可以看出，“民主”一词在辛亥革命之后明显产生了正向的迁移，这代表“民主”一词在辛亥革命之后开始接近了革命的内涵，且这一变化是显著的。但在袁世凯上台之后，“民主”和“革命”之间的相似度又回归到 0 的周围，并且“民主”一词词义的变化不显著。这说明“革命”一词的民主内涵很快消失。“革命”中包含的“自由”的内涵也有相似的变化。但值得注意的是，辛亥革命带来的对“平等”的强调保留了下来。



## 2. 改革内涵的演变

### 2.1. 改革的语义变迁

通过动态词向量模型的计算，我们追踪了不同时期与“改革”最相关的词汇，并对比了与“改革”词义相近的词汇。结果如下表所示。我们观察到，在 1911 年以前，改革的焦点主要集中在政府行政体系的现代化上，涵盖了内阁制度的引入、货币制度的改革、边疆事务的管理，以及司法体系的独立化。这反映了晚清政府在国内外国压力下，试图通过政治和行政改革来应对挑战。进入 1910 年代，改革的重心转向经济领域，尤其是财政和税收体系的改革。到了 1920 年代，改革的方向则转向法律体系和工业领域，特别是矿业的开发和技术引进，显示出国民政府在巩固政权的同时，致力于推动法制现代化和经济工业化。而在 1930 年代，改革的重点进一步扩展到更广泛的政治和社会制度，如省级行政体制、军事组织、政府机构结构，以及教育政策的改革，这可能与国民政府在全国范围内进一步巩固政权，以及推动国家统一和现代化的努力有关。

在晚清时期，“革新”一词与“救国”、“满清”和“留日学生”高度相关，这不仅与当时的历史背景紧密相连，也反映了当时中国社会的紧迫感和危机意识。许多知识分子和政治活动家坚信，只有通过根本性的革新，才能从根本上拯救国家。这种思想在辛亥革命前夕尤为显著。北洋军阀统治时期，“革新”一词的内涵扩展到了更广泛的社会政治领域，涉及政治经济的自主发展、宗教改革、民族独立运动等，反映了北洋军阀时期国内外政治经济形势的复杂性和多元化的社会思潮。到了国民政府时期，革新的重点转向了具体的经济和技术领域，如五年计划的制定、电信技术的发展、商业法律的完善、语言文字的改革和人才培养，这与国民政府推动经济发展和社会现代化的政策密切相关。

与此同时，“改良”一词的相关词汇也在不断变化。在晚清时期，改良主要聚焦于社会和教育领域，例如家庭管理、传统教育体系（私塾）的改革，以及社会问题（如烟膏问题）的解决。同时，对军事征兵制度和刑罚体系（监狱）的改良也体现了晚清政府在国内外压力下的应对策略。北洋军阀时期，改良运动的主要目标转向了经济领域，与“改革”的相关词汇产生了一致性。而在 1920 年代以后，经济领域的改良运动更加具体化，紧密联系着实业的发展。

这一时期的社会变革显示了从传统社会向现代国家的转型过程。早期更多关注政治和行政体系的改革，逐渐转向经济自立和工业化，最后涉及到更广泛的社会和政治制度的改革。这一过程不仅反映了中国社会对内外压力的应对，也体现了对现代化和国家建设的追求。同时，革新和改良的变化也揭示了社会思潮的多元化和对具体实践的重视。

	改革 Reform	革新 Renovation	改良 Improvement
1907	行政，内阁会议，币制，边务，司法独立	帝，满清，部落，留日学生，救国	家政，私塾，教育，烟膏，征兵，习艺所，监狱
1911	贵族，移民，议院，国会，财政	—	监狱，国货，农务，体育，研究所，赛会
1913-1919	币制，盐政，关税，经济，税制	—	盐政，债务，农业，工商部，币制
1920-1927	矿业，规程，组织法，法权，法规	政治经济，自立，宗教，佛教，独立运动，实地调查	植棉，蚕桑，人才，材料，道德，习惯，手工，蚕丝，造纸，国煤，劳工，饼干，罐头食品
1928-1929	系统，制度，技术，禁娼，规程，度量衡	社会教育，设施，学术，公共卫生，军事训练	
1932	法令，设施，内政部，官制，课税	宗教，侨务，文教部，基金会，陈果夫	蚕桑，工业，产业，国际贸易，度量衡，茶业，医药
1934	省制，制度，机构，军制，政制，教育方针	五年计划，电信，商法，语文，人材	实业部，农具，棉统会，蚕桑，茶叶，国煤

而在 1920 年代以后对实业改良的重视也能部分地说明上一节商人和工人在这一时期与“革命”的关联性减弱。1920 年代，随着“改良”一词与“实业”紧密

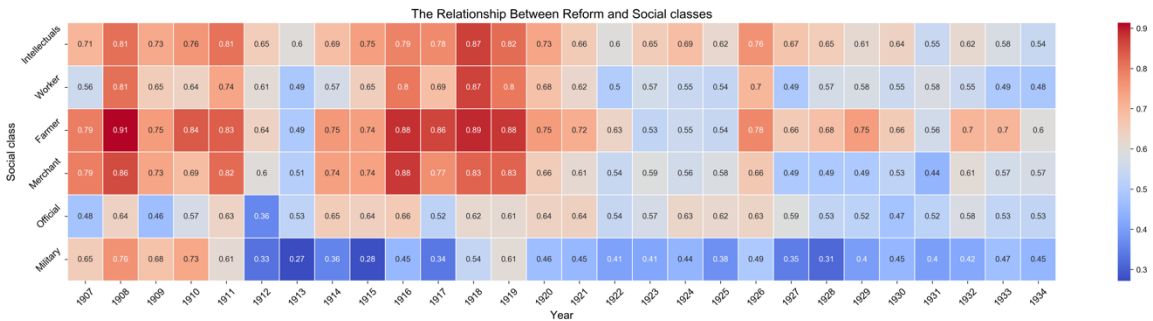


联系起来，中国商人阶层可能更多地将注意力转向了经济发展和商业活动，而不是政治革命。在这个时期，中国的经济民族主义和自力更生的理念开始流行，许多商人开始支持通过发展国内工业和商业来实现国家的现代化和独立，而不是通过政治革命。

此外，国民党政府在北伐后推行的一系列政策，包括建立更加稳定的政府、促进国内工业和实业发展的措施，可能让商人阶层认为他们可以在现有政治体制内实现自己的经济目标。因此，他们可能越来越多地与政府合作，而不是与革命力量合作。

2.2 谁在改革中受益？

我们同样计算了“改革”以及“改革”的近义词与不同社会阶层之间的相关度，希望借此分析哪个阶层主导了改革，或者改革与哪个阶层的利益最相关。



从热力图结果中可以看出，知识分子始终与改革密切相关。从晚清到民国，知识分子积极参与政治讨论，推动改革议程，如维新运动、五四运动等，通过推动教育、政治和文化改革实现国家的现代化。

农民阶层与改革的相关度变化同样值得讨论，从最近邻结果上看，1920 年代出现了农业技术和农具的改良。这一时期“改革”和农民阶层的相关性也开始升高。这说明随着中国的经济发展，特别是城市化和工业化的推进，对农业产品的需求增加，这可能促使更多关注于农业生产效率和技术的改进。蚕桑等产业的发展不仅提高了农民的收入，也可能激发了对更广泛的社会和经济改革的兴趣。

对比这一时期农民阶层与革命之间较弱的相关度，这能够部分地说明农民阶层对于改革和革命的不同态度和需求。他们更关注与自身生产和生活直接相关的改革议题，农民可能更倾向于实际可行的改善措施，而不是抽象的政治理念。

二、国家与国民

1. 国家认同的构建

1.1. 中国的民族主义

中国民族主义的发展轨迹呈现出多元化的特征。在戊戌变法期间，民族议题主要集中在统治阶级的满族与被治理的汉族之间存在的平等问题上（Zarrow, 2012）。然而，随着辛亥革命的胜利，民族主义的焦点从反满及追求满汉平等逐渐转变为追求民族的独立和自我振兴。由于自然语言处理技术的局限，古汉语词汇与

现代汉语词汇在同一向量空间内的融合分析存在困难。因此，在本研究中，我们专注于分析现代汉语，即白话文词向量空间。在探讨“民族主义”与“种族”概念时，我们不得不将《天演论》以及戊戌变法时期关于“民族”和“种族”的文言文定义和论述排除在外。

另外，考虑到满汉问题在中国民族主义发展中的地位，除了“民族”和“种族”以外，我们同样分析了“满人”“满蒙”和“汉人”在语义空间中的变迁。和“民族主义”相关的词汇最近邻在下表中展示。

	民族	种族	满人	满蒙	汉人
1907	-----	罪, 危迫, 兵权, 亡国, 罪人, 不足	-----	国权, 铜像, 纪念	-----
1910	州牧, 外宾, 民主, 官职, 兵官, 治藏	-----	抵御, 扬子江, 杀戮, 轰击, 炮声	同盟, 边界, 争端, 退让, 谈判	痛恨, 清帝, 宗室, 权力
1912	民生, 四万万, 国亡, 社会主义, 救国会, 外患	开辟, 种烟, 牧师, 印铸局, 充军	豫备, 铁良, 上谕, 清国, 君主立宪, 入宫	和解, 缔结, 侵略, 召回, 德法, 英政府	比利亚, 罗马, 哥萨克, 哈尔滨, 维也纳, 华兵, 蒙兵
1918	京奉路, 铜产, 台湾银行, 食物, 银价, 收归国有, 资本家	-----	-----	-----	-----
1922	少数民族, 英俄, 英法德, 鞑靼, 巴伐利亚, 油矿, 英海军	统一党, 国体, 公道, 黠武, 主教, 志愿	-----	国防, 海陆军, 日本政府, 社会主义, 自救, 列国, 亲善	-----
1926	亚洲, 国民外交, 宪法, 省宪决议案, 党部, 制宪, 国民大会	-----	-----	最后通牒, 野心, 俄蒙, 内政外交, 国防	-----
1930	肺结核, 疗法, 墨索里尼, 国民经济, 都市, 全世界, 南美	-----	-----	满洲, 并吞, 领土, 军阀, 帝国主义, 国际公法	-----
1934	智识, 宗教, 同志, 精神, 文学, 儿童教育, 心理, 语文, 教学	生育, 劳动者, 日耳曼, 潮流, 生殖, 优生	水田, 禁售, 营盘, 茧丝, 鲜农, 火患, 柞蚕, 商户, 业绩	采金, 制纸, 海产, 资本金, 织品	-----

我们可以看到在 1912 年，辛亥革命刚刚成功，推翻了清朝统治，建立了中华民国。这一时期的民族主义主要围绕着反对帝制、推崇共和和民族独立的理念。救亡图存（即救国图存）成为当时民族主义的核心议题，反映了中国人民对于国家存亡的紧迫感和民族自救的意识。

经过辛亥革命后，中国社会逐渐意识到单纯的政治改革不足以实现国家的强大，经济和实业的发展同样重要。这一时期的民族主义开始强调经济自强，提倡“富国强兵”的理念，认为经济独立是实现民族独立的基础。因此在 1918 年，“民族主义”开始和经济发展紧密联系在一起。

1920 年代，中国深受列强侵略和内部战乱影响，民族主义在这一时期更多地关注国际政治格局。五四运动后，民族主义与反帝、争取国际平等地位的思想紧密相连。中国开始寻求在国际舞台上维护自身权益，民族主义与国际关系的联系因此加强。

1926 年，民族主义开始和政治架构产生关联，“国民外交”、“宪法”、“省宪”、“制宪”和“国民大会”等词汇与国民政府的政治架构和活动有关。这一时期正是国民政府试图在中国建立更加中央集权的现代国家的时期，这包括制定宪法、推行国民外交政策，以及召开国民大会等活动。

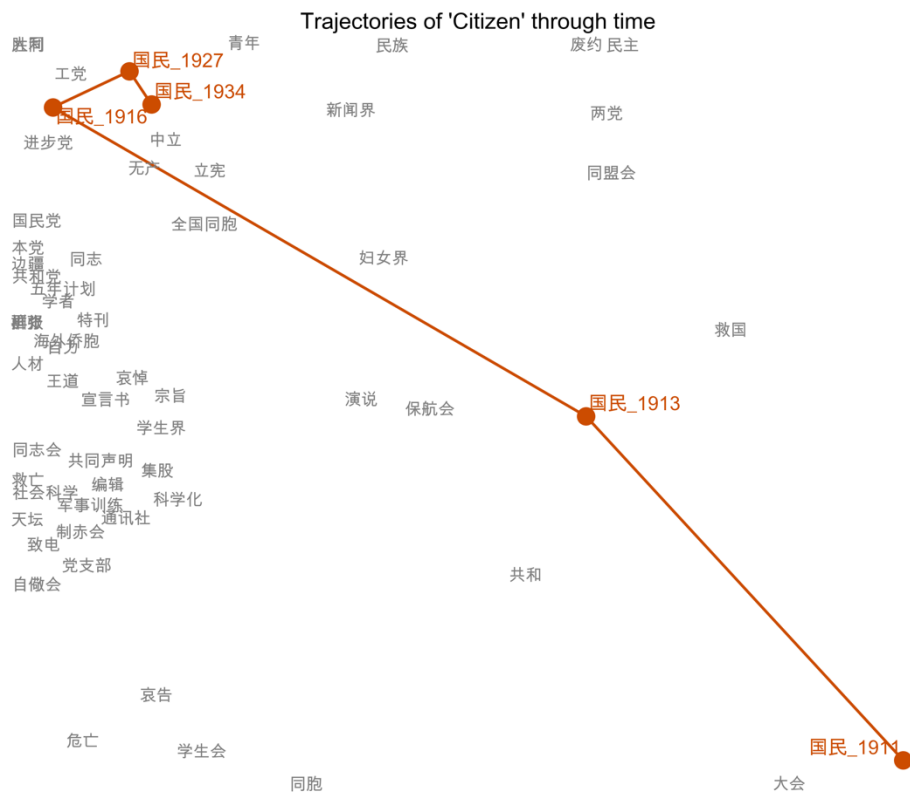
到了 1930 年代，随着国内政治的相对稳定和抗日战争的临近，民族主义开始更多关注内部建设，特别是教育和科学的发展。这一时期，民族主义强调通过教育和科学的进步来提升国家整体实力，认为这是抵御外来侵略和实现民族复兴的关键。另外值得注意的是，在 1930 年代，“民族主义”相关词汇中开始出现了与法西斯相关的词如“墨索里尼”“日耳曼”，这种世界意识形态的流动可能似的中国的民族主义在某些方面采纳或受到了当时法西斯主义强调的某些主张，如民族精英主义、民族使命感、种族优越感等的影响。

从对满汉民族矛盾的角度上看，似乎这个讨论在辛亥革命之后很快的消失在大众舆论之中，在 1912 年以后对“满人”和“汉人”的讨论基本消失，在词向量空间中也没有相应的词汇出现。而 1920 年代到 1930 年代对“满蒙”的讨论都聚焦在国际关系的范畴中。民族主义在辛亥革命之后很快由对国内民族平等的诉求转向了面对外部威胁时对内部团结的塑造和激励。

## 2. 国民的角色与地位

### 2.1. 国民身份的构建

我们计算了“公民”、“国民”、“市民”这些词在词向量空间中的最近邻。通过分析目标词在不同时间点的最近邻，我们可以观察到与目标词关系最密切的词汇是如何随时间发生变化的。这不仅反映了目标词本身语义的演变，也映射了它与语言网络中其他词汇的动态关系。



我们可以看到，“国民”一词在 1916 年以后在空间上聚集，这代表“国民”一词的语义在 1916 年以后相对保持稳定。“公民”一词在空间上发生了偏移：在 1910 年代“公民”与政治参与有关，但在 1926 年以后开始向“公民教育”相关的概念偏移。“公民”一词在向量空间中的这种变化，从与“选举”等政治活动相关联，到后来与“教育”、“教育部”、“大学”等教育领域的词汇相关联，反映了该词语义上的演进和社会焦点的转移。



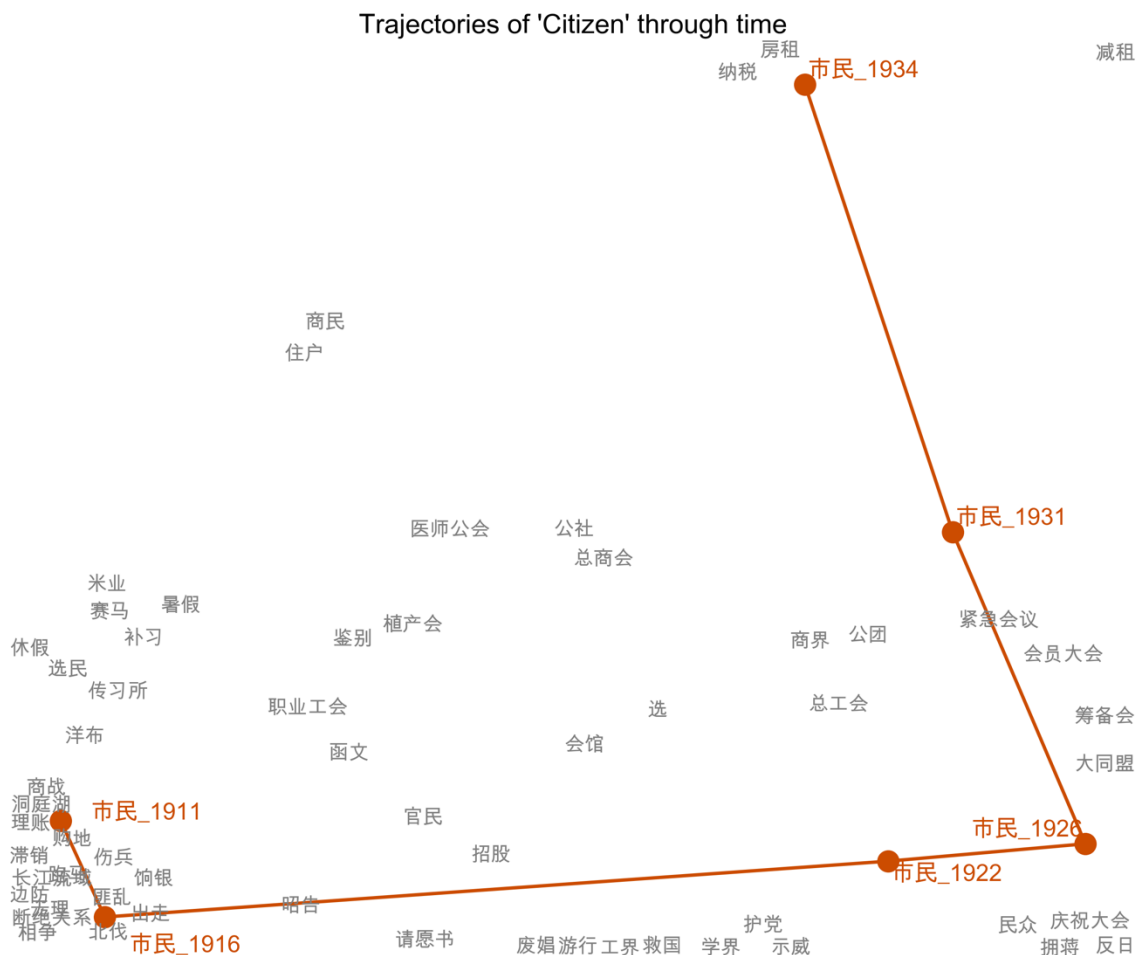
*Post\_Treatment*的回归结果，当 $\beta_3$ 显著时，代表着在特定历史时期以后，在上下文出现了“公民”的语段中，我们关注的目标词词义出现了显著地变化。如果我们观察到 $\beta_3$ 显著地正向移动，则代表 Y 与“公民”在语义上更为接近。

下表给出了回归的结果，结果表明，在袁世凯称帝之后，社会舆论对“公民”的讨论更多地与“选举”这一政治权利联系在一起。并且随着时间的流逝，“公民”和“选举”在词义上更为接近。在蒋介石掌权之后，“公民”和“教育”在语义空间上更为接近。

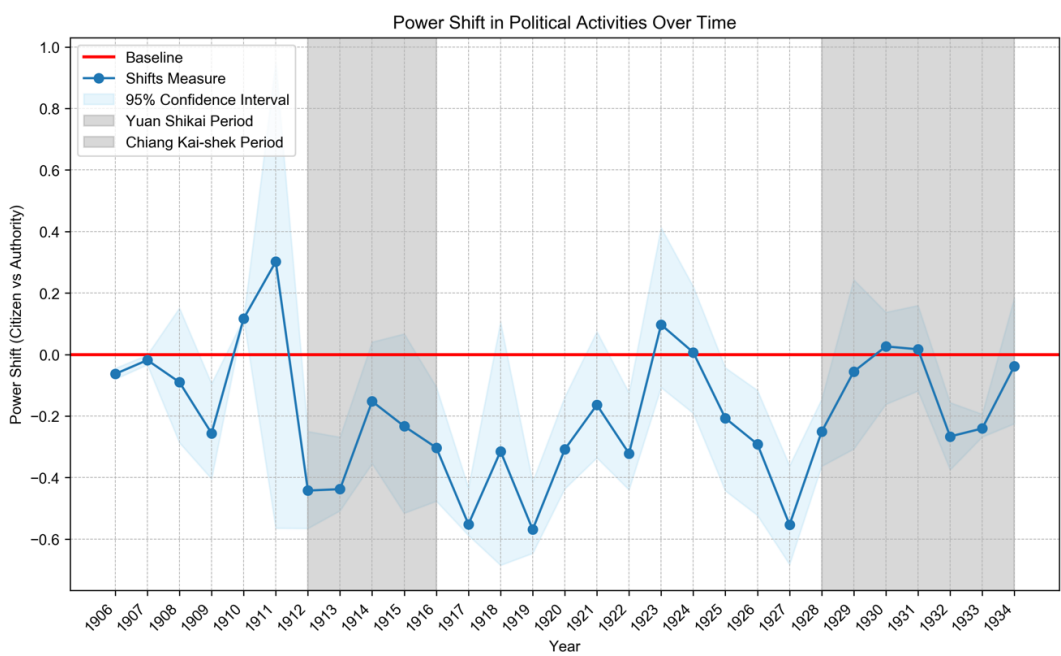
	选举	权利	教育	道德
Baseline (1912/1/1)	0.0008	1.07	0.00237	1.703
	(3.382)	(2.173)	2.445	4.291
Treat1 (1913/7/1)	0.002	0.05	0.00107	1.724
	(4.303)	(1.688)	(3.34)	(3.632)
Treat2 (1915/12/31)	0.002***	1.051	0.0006	1.675
	(0.0003)	(2.056)	(0.698)	2.959
Treat3 (1927/2/1)	0.003***	0.006	0.003***	-
	(0.0001)	(0.619)	(0.0003)	-
Documents	15749	23680	978	962

“市民”一词在向量空间中的变化，从与“洋布”、“米业”、“赛马”等词汇相近，转向与“游行”、“集会”、“同盟”、“会议”等词汇相近，代表了该词语义上的显著转变。这种转变反映了“市民”一词所承载的社会和文化含义的演进。在早期，“市民”一词与“洋布”、“米业”、“赛马”等词汇的接近性可能指向了市民生活的日常经济活动和娱乐休闲方面，这些词汇通常与商业交易、物质消费以及社会生活的某些方面相关联。这表明在这一时期，“市民”一词在语言使用中更多地与经济活动和社会生活的普通层面联系在一起。然而，随着时间的推移，“市民”一词的最近邻转向了“游行”、“集会”、“同盟”、“会议”等，这些词汇与公民参与、政治活动和社会运动紧密相关。这表明“市民”一词的含义开始与公民的政治权利、社会责任以及积极参与社会公共事务的概念联系起来。这样的语义演变可能与社会政治环境的变化有关，如民主化进程、公民权利意识的提升、以及社会运动的兴起。





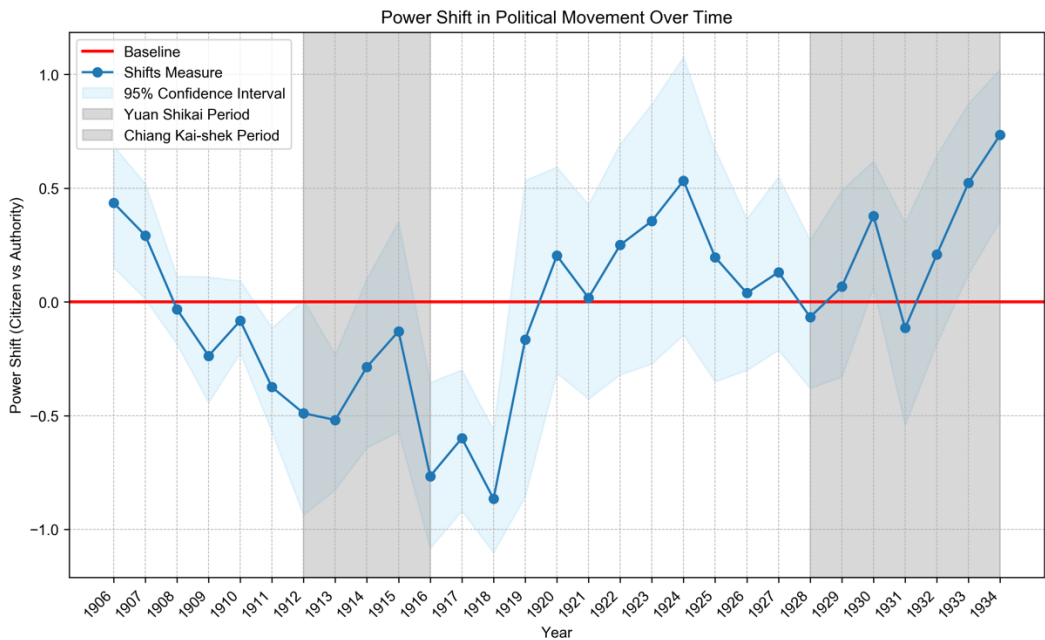
我们通过图表形象地展示了这一变化过程，其中阴影区域代表了参与者角色偏移的 95%置信区间，为我们的结论提供了直观的证据。与代表性群体向量距离差异为正，表明在这一时期，政治活动的主要参与者更倾向于普通公民；而距离差异为负则意味着主要参与者更接近当权者。这种分析不仅揭示了公民与当权者在政治活动中的角色变化，而且还反映了公民身份认同和国家现代性概念的演进。通过这种方法，我们发现在中华民国成立之初，普通公民在国庆等政治活动中的参与度较高，这表明了公民身份的积极塑造和对国家现代性的共鸣。然而，随着时间的推移，公民对国家庆典的热情有所降温，主要参与者逐渐回归到了当权者阵营。1920 年代初期，公民的参与热情虽有短暂回升，但随着国共矛盾的加剧，这种热情再次降低。



注：点向正向移动代表这一时期国家庆典活动的主要参与者为普通市民，点向负向移动代表这一时期国家庆典活动的主要参与者为当权者

为了进一步阐明政治活动与公民身份塑造之间的关系，我们选择了一系列政治运动来构建一个考察公民政治参与度的中性词组，其中包括诸如‘革命’、‘游行’、‘政治参与’、‘抗议’、‘罢工’、‘反抗’、‘示威’等词汇。研究结果显示，在 20 世纪的第一个十年，政治活动与社会变革更加靠近官员群体。这可能是因为这个时期的社会运动更加强调知识分子和思想家的参与，他们通常更接近政府和官员群体。而在 1920 年以后，政治活动与社会变革更加接近市民群体，这可能反映了公民社会的发展，市民和普通人更加积极地参与政治活动和社会运动。此外，“市民”一词的最近邻结果同样表明，在 1920 年以后，中国的公民社会逐渐

壮大，“市民”一词在空间上开始接近工会、学生组织、妇女联合会等组织。这些组织和运动推动了社会的变革，增强了市民在政治上的影响力。



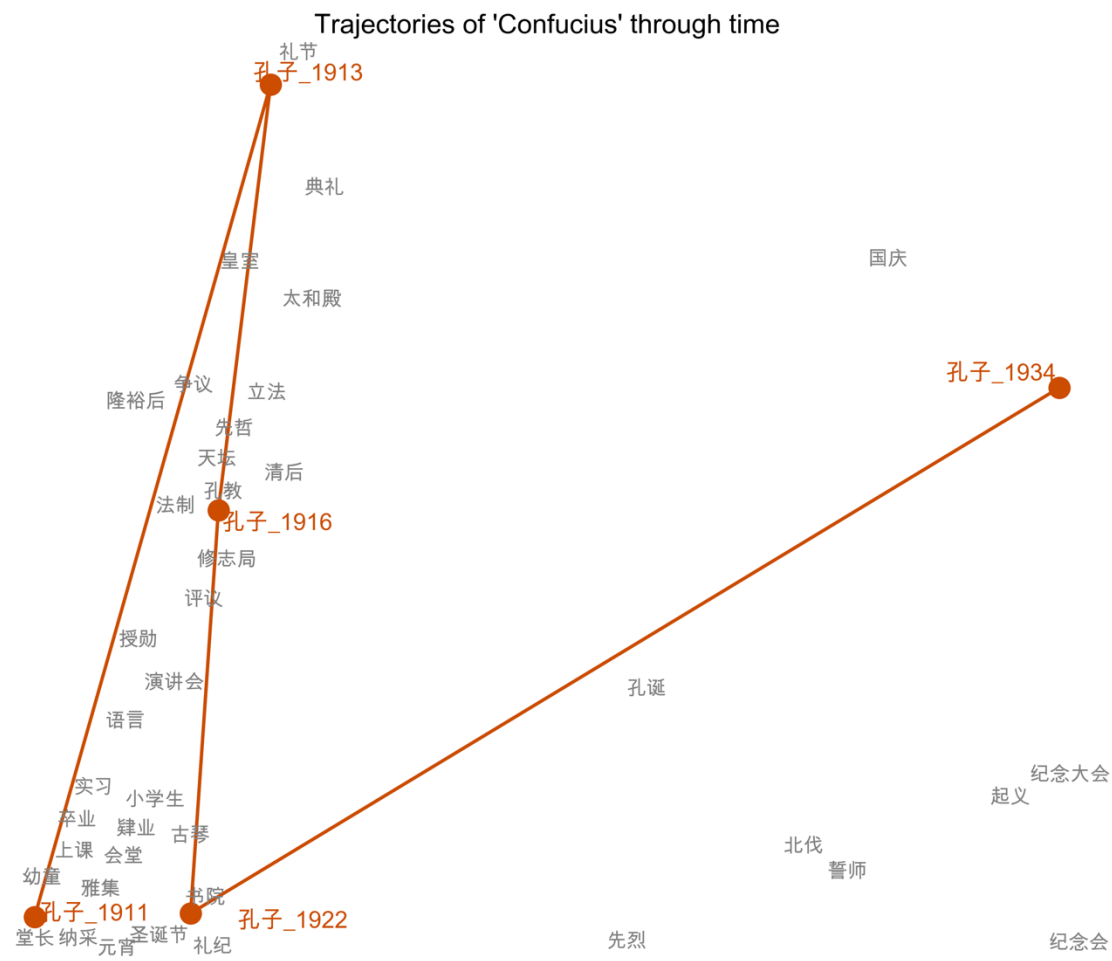
注：点向正向移动代表这一时期政治活动的主要参与者为普通市民，点向负向移动代表这一时期政治活动的主要参与者为官员群体

三、文化与宗教

1. “孔子”与“孔教”的塑造

1.1 语义的演变

这一部分我们同样使用了词向量的最近邻和降维计算。从“孔子”一词的二维词向量动态轨迹上我们可以看到1911年“孔子”一词在语义上与“教育”紧密相关，但随着袁世凯上台，“孔子”的含义开始发生偏移。在1913年至1916年间，“孔子”一词更多地与“天坛”、“祭典”和“授勋”等概念联系在一起。而在20世纪20年代初，“孔子”一词的含义又回归到了“教育”。到了30年代，“孔子”一词又被赋予了与国民党政治理念相关的革命性内涵，在语义空间中与“起义”、“北伐”和“革命先烈”等概念联系起来。而“孔教”一词则逐渐开始与蒋介石倡导的国家道德相关。



鉴于代表孔子和儒家学说的词汇随时间也在变化，我们基于词向量之间的余弦相似度，计算了“孔教”和“儒学”的最近邻词汇，并在表格中展示了结果。与词向量投影结果相似，1913 年至 1916 年间，“孔教”一词被创造出来，将孔子学说塑造为“国教”的说法开始兴起。在这一时期，“孔子”和“孔教”开始与国家祭典和法治概念紧密关联。但到了 20 年代初，“孔子”和“孔教”的含义又回归到教育和传统领域。1928 年之后，“孔子”开始与国民党的革命和道德理念相联系。到了 1930 年代，“儒学”一词出现，并成为“新文学运动”的一部分。通过最近邻分析，我们不仅能够追踪“孔子”一词的语义变迁，还能洞察与之相关的词汇如何随着历史背景的变化而演进。

表 1:与“孔子”相关词的最近邻

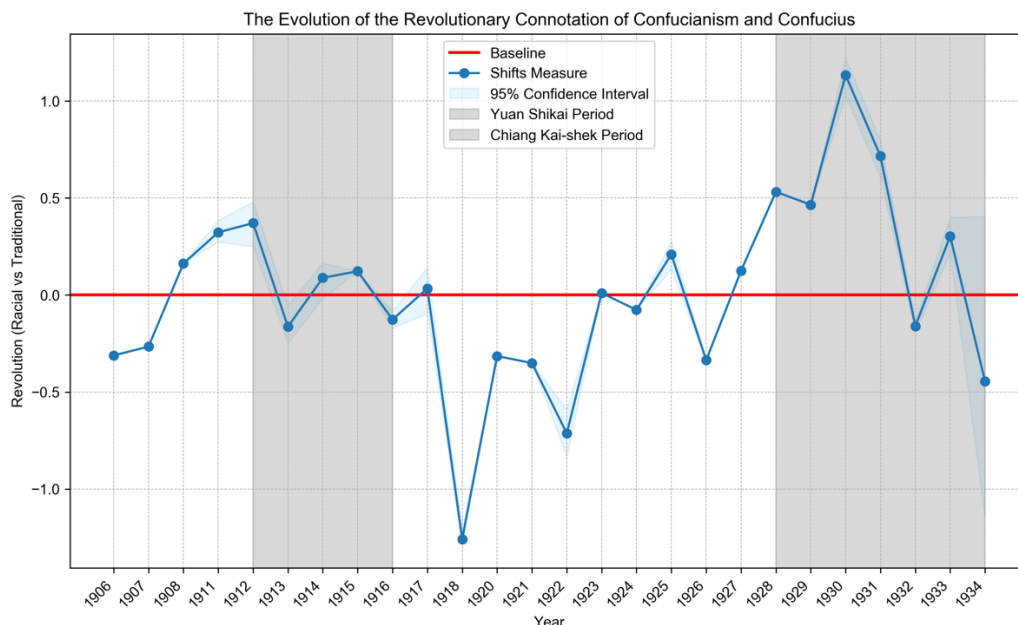
Year	“孔子”	“孔教”	“儒学”
1911	肄业，上课，幼 童，小学生	—	—

1913	授勋，皇室，天坛	国教，救国，教会，立宪	—
1916	孔教，立法	立法，宪政，评议会	—
1922-1926	书院，雅集，古琴，礼记，学子，学馆	耆绅，公牒，家族，故绅，大学，课业	—
1928-1934	纪念大会，北伐，革命先烈，紫金山，遗像，黄花岗	精诚团结，立国，万民，美满，早婚	新文学运动

### 1.2 孔子革命性内涵的演进

孔子和孔教在革命内涵上的演进是一个较为复杂的过程，这一过程不仅体现语义变化上，而且在历史事件和国家仪式中有着具体的体现。在清朝末年，孔子的祭祀被提升为大祭，这一变化反映了改革派如康有为等人的改革愿景，他们视孔子为中华民族的象征和中华文化的精髓。清朝政府希望通过提升孔子的地位来维系国家的统一，尤其是在科举制度被废除后。这一时期的孔教不再仅仅是传统和权威的象征，而是开始与民族主义和现代爱国主义的概念相结合。清朝对孔子的新提升是试图加强国家认同和人民与君主制之间的联系，但这牺牲了皇帝的超然地位，根据康有为的观点，皇帝只是祭孔活动中的参与者之一。在这一时期，孔子的现代性体现在对平等的反复强调上，尤其是皇权与民权之间的平等。

进入民国时期，孔子的角色继续演变。政治领袖如袁世凯试图利用孔子来构建新的国家认同。通过将孔子与民族主义和民主思想相结合，儒学的某些元素被重新诠释，以适应推动社会变革的力量。为了证明孔教中革命性内涵的演进，我们选取了两组与革命演变相关的词汇列表，一组代表着晚清康梁一派的、早期的革命理念，词汇列表由“平等”“公平”“自由”“博爱”“仁义”等构成，而另一组代表民国时期较为激进的革命理念，词汇列表由“革命”“起义”“烈士”“义士”等构成。我们同样计算了代表孔子的一组中性词（“孔子”“孔教”“儒教”“儒家”等）与两组不同的革命理念之间的欧几里得距离差，得到了衡量革命性内涵演进的度量指标。



注：点向正向移动代表这一时期“孔子”的革命性内涵更为激进，点向负向移动代表这一时期“孔子”的革命性内涵更接近早期革命内涵。

词向量模型的分析进一步证实了与孔教相关的词汇在语义空间中的革命性概念演变。在袁世凯试图复辟帝制的那段时期，进步的知识分子们开始将他的专制倾向归咎于儒家主义的束缚。在这一时期，围绕孔子及儒家学说的讨论逐渐聚焦于对等级伦理的批判，以及探讨儒家主义是否能够促进公民权利的彻底平等。进入民国的第二个十年，孔子再次被赋予了全新的革命形象，对儒家学说的解读与起义等革命行为关联性增强，与平等的公民权利之间的关联渐弱。特别是在1930年之后，与孔子相关的祭祀活动更加强调对革命烈士的纪念，尤其是对黄花岗起义烈士的崇敬，这反映了孔教在革命内涵上的演进。

### 1.3 孔子与儒家的宗教化倾向

在1912年，《时报》的一篇社论中提出，正是儒家思想的深远影响使得中国这个国家得以历经磨难而长存：面对外来侵略的屡屡挑战，中国始终能够再次振兴。20世纪初期，儒教的倡导者们认为，儒家教义中关于平等与公正的理念，是现代民主的最佳体现，而过去的“天道”则被封建皇帝所独占。到了1910年代，关于将儒家教义转化为一种宗教信仰的讨论逐渐兴起。陈焕章将儒家视为西方宗教的对等物，因为它独特地将道德与政治结合在一起，并将孔子置于类似于神圣的地位。即便是那些反对将儒教正式确立为国家宗教的人士，他们大多也是儒家思想的拥护者，这表明他们的反对并非针对儒教本身，而是反对其宗教化的正式过程。这种态度反映了儒教作为一种道德秩序与其作为潜在宗教体系之间的内在张力。

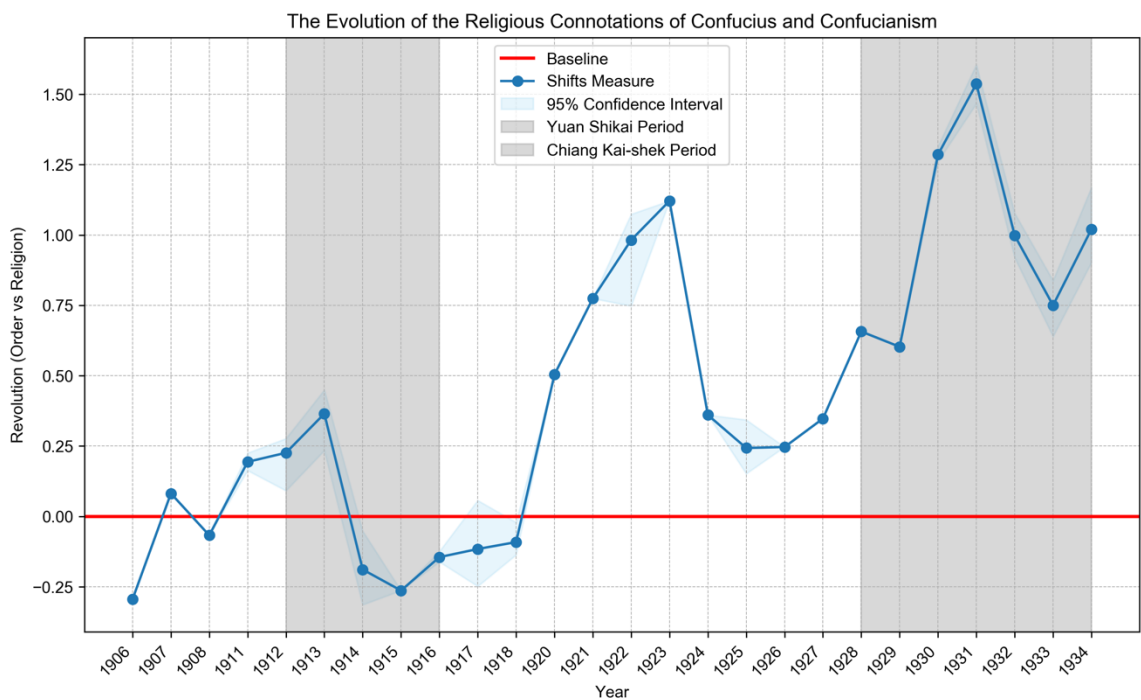
1934年，蒋介石决定推行新生活运动，该运动的核心理念与儒家学说紧密相连，尤其是在强调个人道德修养和社会责任方面。儒家学说倡导“修身、齐家、治



国、平天下”的连贯理念，即从个人的道德修养出发，逐步扩展至家庭、国家，乃至全世界的和谐治理。蒋介石的新生活运动正是试图将这种儒家的道德修养理念应用于国家治理和社会改革之中，以实现国家的稳定与繁荣。在 20 世纪 30 年代对孔子的讨论中，儒家思想的核心再次回归到了传统道德秩序的框架之内。

我们通过一组代表传统道德秩序的关键词汇（如“生活秩序”、“道德”、“礼法”、“纲常”）来衡量孔子的世俗道德倾向，同时通过一组代表宗教的关键词汇（如“寺庙”、“天主教”、“佛教”、“宗教”）来衡量孔子的宗教化倾向。我们还测量了代表孔子的词组与这两组关键词汇之间的欧几里得距离差异。结果显示，1910 年至 1914 年间的儒教讨论中出现了明显的宗教化倾向，而在蒋介石的统治时期，孔子和儒家学说则被赋予了更为强烈的道德秩序特征。

综上所述，孔子和儒教在革命内涵上的演进是文化、社会和政治互动的综合结果。通过量化分析方法，我们能够更客观地理解儒教如何适应并被重新解释，以应对社会变革的需求。这种演进在词向量模型中的显现，即与儒教相关的词汇与革命性词汇之间距离的变化，为我们提供了孔教如何随时间演变并影响社会和政治革命角色的洞见。



注：点向正向移动代表此时“孔子”与道德和秩序更相关，点向负向移动代表此时“孔子”与宗教更相关。