·社会学研究·

DOI:10.20004/j.cnki.ujn.2024.05.011

村庄价值观的共同基调与地方风俗

——基于万级村庄村规民约文本的大数据分析

吕 鹏[1],刘金龙[2]

(1.中国社会科学院 社会学研究所,北京 100732;2.中央民族大学 民族学与社会学学院,北京 100081)

[摘 要]村规民约在中国乡村治理中扮演着成风化俗的角色,是推进农村现代化尤其是精神文明建设的重要手段,是日用的共同价值理念的文本呈现。现有研究受限于小样本和传统分析方法,难以全面揭示其内在的多样性。利用计算社会科学倡导的无监督的文本聚类模型和有监督的文本分类模型,对某数字治理平台上覆盖全国范围内将近1.5万个村庄的村规民约文本进行深入分析,挖掘其中的治理主题、内容分布及其地方性特征。研究为理解村规民约在中国乡村治理中的作用提供了新的视角,也为国家与地方治理规范的互动提供了新的理解。

[关键词]村规民约;乡村治理;主题分类;预训练模型;计算社会科学

[中图分类号]C912.69 [文献标识码]A [文章编号]1671-3842(2024)05-0111-11

一、既往研究与问题提出

村规民约作为村民进行自我管理、自我服务、自我教育、自我监督的行为规范,是完善自治、德治、法治相结合的现代基层治理体系的重要形式。习近平总书记指出,“推进农村现代化,不仅物质生活要富裕,精神生活也要富足。农村精神文明建设要同传承优秀农耕文化结合起来,同农民群众日用而不觉的共同价值理念结合起来,弘扬敦亲睦邻、守望相助、诚信重礼的乡风民风。”①。长久以来,村规民约是中国乡村自治的重要依据。在广泛的基层治理实践中,村规民约逐渐成为地方治理经验与现代化国家治理理念、治理能力紧密结合的成果②。

学者们广泛使用“国家—社会”的视角来分析村规民约。该视角认为国家法律与村规民约体现着正式制度与非正式制度的关系③。由于国家法律的一般性和抽象性使得其很难在基层实践中得到贯彻和运用,因此在基层国家正式制度一般变通为非正式制度形式运行④。村规民约曾经被视为一种地方性非正式制度⑤,其内容更加注重本乡本土具体的历史、现实、风俗、特点,在基层治理中

[基金项目]中国社会科学院实验室孵化专项资助项目“社会计算与数字社会前沿研究”(项目编号:2024SY-FH010)。

[作者简介]吕鹏,中国社会科学院社会学研究所,中国社会科学院大学数字中国研究中心教授、博士生导师,邮箱:lv-peng@cass.org.cn;刘金龙,中央民族大学民族学与社会学学院博士研究生。

①习近平:《加快建设农业强国 推进农业农村现代化》,《求是》,2023年第6期。

②赖先进:《发挥村规民约在社会治理中的耦合协同效应和作用》,《科学社会主义》,2017年第2期。

③杨开道:《中国乡约制度》,北京:商务印书馆,2017年版,第185页。

④孙立平,郭于华:《“软硬兼施”:正式权力非正式运作的过程分析——华北B镇收粮的个案研究》,《清华社会学评论》(特辑1),厦门:鹭江出版社,2000年版,第21页。

⑤李喜英,高维,申翠叶:《社会资本视角下创新传统乡规民约的思考》,《广西大学学报》(哲学社会科学版),2020年第2期。

发挥着重要的协调和组织作用,村规民约的内容包含着村庄“成员”的资格、享受权利的条件、个体和整体关系协调[①]。例如,卢瑶玥等人将村规民约内容与农村养老福利结合分析发现,村规民约的联结度和可及性越强,村庄内部养老福利供给越强[②]。村规民约以书面签约的正式制度形式,从生活、生产的方方面面介入村民的日常生活,在村庄内部发挥着重要的整合功能[③]。

随着国家政权建设的不断深入,村规民约的内容已经不止是出于乡村治理的本地需要[④],村规民约也开始“现代化转型”[⑤]。不少研究基于内容分析发现当前村规民约内容的标准化、法治化程度不断得到提升[⑥],民间秩序正呈现出民主化、制度化以及与国家秩序、法律秩序协同的趋势[⑦]。郎友兴等人利用词频统计方法对某个村庄十个版本的村规民约分析发现,随着经济社会的发展,村规民约中话语从义务责任制论的治理开始转向维护权利论的治理,从整合控制走向重视主体性的倡导[⑧]。从原因来看,有学者认为这是由于中国社会的巨大结构变迁,乡村社会正呈现出“结构混乱”的状态,乡村法律实践的场景和逻辑也随之发生变迁[⑨]。村规民约中地方性不断被削弱,村规民约在与国家法律的博弈中逐渐退出[⑩],在表现形式上会与国家正式法律制度愈发趋近,取而代之的是国家的意志,有学者将其概括为村规民约的“国家化”[⑪]。也有研究认为,大量规范化、制度化村规民约的涌现,并非都是村民自治自我发展的结果,其中有相当一部分是政府为实现地方治理而在背后推动的结果[⑫]。村规民约不仅将国家权力进行地方化,同时也是国家法律规则体系的重要补充[⑬]。周家明等人认为村规民约中具备了制度的三大基础要素:规制性基础要素、规范性基础要素、文化认知性基础要素,这三大要素外化表现为惩戒监督机制、价值导向机制、传递内化机制,从而对村庄治理产生影响[⑭]。

乡村治理的实践不仅是权力运作的场域,同时也是文化价值重构与社会整合的过程。虽然“国家—社会”视角有助于我们理解当前村规民约中“国家化”的趋势,但是我们也不能否认在乡村治理中的地方性知识的作用与价值,如果简单将村规民约看作国家法律法规、制度规则的翻版,则无

①张静:《基层政权:乡村制度诸问题》,北京:社会科学文献出版社,2018年修订版,第86页。

②卢瑶玥,覃诚,方向明:《村规民约对农村养老福利的作用机制分析——基于浙江省衢州市28个村的观察》,《中国农村观察》,2023年第2期。

③周怡:《共同体整合的制度环境:惯习与村规民约——H村个案研究》,《社会学研究》,2005年第6期。

④谢晖:《当代中国的乡民社会、乡规民约及其遭遇》,《东岳论丛》,2004年第4期。

⑤丁立磊:《传统乡规民约何以实现现代化转型》,《人民论坛》,2020年第14期。

⑥陈荣卓,李梦兰,马豪豪:《国家治理视角下的村规民约:现代转型与发展进路——基于“2019年全国优秀村规民约”的案例分析》,《中国农村观察》,2021年第5期。

⑦李朝晖:《民间秩序的重建——从乡规民约的变迁中透视民间秩序与国家秩序的协同趋势》,《学术研究》,2001年第12期。

⑧郎友兴,陈文文,薛晓婧:《村规民约的文本和背后的治理议题及演变——以浙江省衢州市上洋村十版“村规民约”为例》,《治理研究》,2022年第5期。

⑨董磊明,陈柏峰,聂良波:《结构混乱与迎法下乡——河南宋村法律实践的解读》,《中国社会科学》,2008年第5期。

⑩姜裕富:《村规民约的效力:道德压制,抑或法律威慑》,《青岛农业大学学报》(社会科学版),2010年第1期。

⑪陈锋,徐悦鑫:《国家力量的双重进入与村规民约的重塑——基于215份优秀村规民约的文本分析》,《学习与实践》,2024年第3期。

⑫张斌:《村寨里的“国家”:基层治理视野下的村规民约——以贵州某地为切入》,《甘肃政法学院学报》,2015年第6期。

⑬高其才:《延续法统:村规民约对固有习惯法的传承——以贵州省锦屏县平秋镇魁胆村为考察对象》,《法学杂志》,2017年第9期。

⑭周家明,刘祖云:《村规民约的内在作用机制研究——基于要素—作用机制的分析框架》,《农业经济问题》,2014年第4期。

助于我们理解乡村治理的复杂性。尤其是考虑到中国不同区域的差异性，也使得我们不仅要关注到村规民约在纵向上随着时代转换带来的差异性，也要关注横向上不同区域之间的差异性。挖掘不同区域之间村规民约主题内容，不仅可以展示不同区域间的差异性，也能更全面了解村规民约中“地方性”。村规民约中各类条款的重要程度，不仅反映了乡村社会治理的实践逻辑，也展示了在国家宏观政策、地方文化传统和村民主体行动之间的复杂互动。

目前已有一些研究对多个村庄的村规民约文本内容进行分析。例如，陈峰等人通过对215个优秀村庄村规民约文本分析发现，国家权力主要通过国家的“村规民约化”和村规民约的“国家化”两种方式进入到乡土社会中，国家力量与地方非正式规范的遭遇中，嵌入地方知识的同时将其纳入到国家的规则体系之下规范化地运行[①]。陈荣卓等人对2019年74个全国优秀村规民约的分析发现，村规民约的规范化、标准化不断突出，治理内容和规则不断与国家趋同[②]。吴晓玲等人对明代以来的村规民约分析发现，中国的村规民约经历了民间自发到半官化再到完全官化、地方性到全国性、自治性到控制性的转变。村规民约的自治属性不断丧失，其内涵和职能也不断为官方所控制，最终沦为官方控制基层的工具[③]。

然而，上述关于村规民约的分析仍有不少值得推进的地方。其中一点就是，这些文章虽然在村规民约的治理功能、治理形式和治理效果等方面做出了有益的探索，然而受限于数据可及性和方法适用性，这些基于村规民约的分析主要局限于单个村庄或小样本村庄的文本词频统计层面。因此，一方面难以实现对村规民约全面、深入的理解，另一方面也无法充分揭示其在不同地区、不同文化背景下的多样性和复杂性，有的还陷入盲人摸象的困境。本文尝试利用计算社会科学的方法，通过大数据分析，挖掘全国范围内不同区域村规民约中治理主题和内容，以及村规民约中地方性色彩。拟回答以下几个问题：首先，当前在全国的村规民约中治理主题主要包含哪些内容，这些主题在村规民约中的分布；其次，不同地区的村规民约是否有显著差异，而这种差异的社会基础又是什么。在理论层面，与既有国家视角的理论预设不同，本文直接对村规民约文本展开分析，以此更加全面展现村规民约中所容纳的内容，挖掘村规民约中的“地方性”。

二、方法与数据

（一）数据来源

本研究数据来源于农业农村部和腾讯公司共同实施的“耕耘者”振兴计划（以下简称“耕耘者”）。“耕耘者”有很多内容，其中一个重中之重，就是推广使用“村级事务管理/服务平台”。简单来说，这个平台是一个微信小程序，为开展乡村治理开发了一系列的数字化工具，通过村规民约、积分规则、评议规则、积分公示等数字化，帮助村庄提升治理效率。在推广的过程中，腾讯与农业农村部、各级地方政府合作，开展了大量的工作，尤其是在四川、广东、新疆、广西等地取得了较大的突破。截至2024年8月15日，耕耘者实际上线村庄65787个，实习村庄31436个，一星村庄10416个，二星村庄12595个，三星村庄7154个，四星村庄2011个，五星村庄2175个[④]。在本研究我们爬取了

①陈锋，徐悦鑫：《国家力量的双重进入与村规民约的重塑——基于215份优秀村规民约的文本分析》，《学习与实践》，2024年第3期。

②陈荣卓，李梦兰，马豪豪：《国家治理视角下的村规民约：现代转型与发展进路——基于“2019年全国优秀村规民约”的案例分析》，《中国农村观察》，2021年第5期。

③吴晓玲，张杨：《论乡规民约的发展及其演变》，《广西社会科学》，2012年第8期。

④村庄星级是根据村庄在数字平台使用状况和活跃度进行评级，村庄星级越高代表数字化平台使用状况越好。其中实习村庄代表村庄仅接入平台，但并未使用。

14954个村庄的村规民约,在进行初步清理后,纳入分析样本14850个村庄[①],涵盖全国31个省直辖市或自治区。其中四川省有8038个村庄,广东省1177个村庄,广西壮族自治区993个村庄,湖北省695个村庄,新疆维吾尔自治区425个村庄,江西省410个村庄,黑龙江省360个村庄。在村庄星级分布上,五星村庄1306个,四星村庄861个,三星村庄3299个,两星村庄4816个,一星村庄1408个,实习村庄3160个。这一体量的数据有助于我们更加全面了解地方性,理解村规民约在不同地区的应用与实践。

(二)研究方法

根据前文分析,需要在有效确定村规民约基本内容的基础上,区分出每个村庄村规民约的具体侧重,最终展开针对全样本的宏观分析。这包括两大任务:首先,需要合理有效地确定村规民约包含哪些内容的文本(或"主题"),建立有关村规民约内容的描述体系,这一步骤带有类型学分析色彩;其次,需要逐份确认不同村规民约中各类文本的具体分布,即某份村规民约中,不同类型文本的篇幅和位置,以及彼此之间的分布关系。

为了完成上述目标,本文在先验知识和专家经验的基础上,混合采用无监督的文本聚类模型和有监督的文本分类模型。

1. 数据标注

文本聚类是一类常用的自然语言处理方法,旨在将相似的文本数据进行归类,并不依赖已有的分类知识[②]。相较于传统的关键词比较或人工分类方法,文本聚类的优势在于可以以统一、稳定的标准,在较短时间内处理大量文本数据,同时无需事先对数据进行分类或标记,也有可能发现未知的类别和模式,但也存在聚类结果不稳定、评判标准缺失以及可解释性较差等潜在劣势。基于文本聚类模型的上述特征,为获得村规民约文本中的主题分布,并进一步推断出村规民约主要内容的内容结构,本文选用常用的隐性迪利克雷(Latent Dirichlet Allocation,LDA)模型作为获得文本类型体系的基础模型。

不可忽略的是,LDA模型和其他文本聚类模型一样,同样可能面临稳定性不佳、可解释性较差等问题,其关键词的解读和主题的归纳也有赖于模型使用者的主观判断。因此,为了得到有关村规民约正确、有意义的分类体系,有必要运用先验知识、专家知识共同构建。腾讯"耕耘者"平台通过与农业农村部相关司局、国内顶级高校和研究机构的合作,聘请了包括作者团队在内的众多专家学者全过程深度参与到平台的论证、开发和评估工作,因此在村庄使用平台、上传村规民约时,平台给予了一套现成的分类体系以供参考。本文利用了这一分类作为村规民约形成的先验知识。同时,本文还基于既有研究和专家意见,对村规民约原始文本和村规民约文本类型体系进行初步解读、评价、修改。经过模型多轮迭代和多次模型结果—人类知识对齐过程,最终可以获得一套稳定的村规民约内容体系。

2. 预训练模型

文本分类是一种将文本数据分配到预定义类别中的过程,它涉及到对大量非结构化的文字信息进行分类,如书籍、文档等,根据其内容分到指定的类别中去。本研究使用深度学习方法中的Bi-directional Encoder Representations from Transformers(BERT)模型进行文本分类。为了实现这一目标,本文利用Pytorch和Transformers库,构建了一个基于预训练模型的文本分类器。该分类器的核心组件包括数据预处理方法、模型设定和训练方法以及模型评估和预测方法。出于控制任务复杂性和算力成本的考虑,本文选用BERT模型作为预训练模型,在BERT基础上进行微调得到上述最

①为了保证分析的一致性,同时考虑到部分村庄村规民约形式采用三字经或顺口溜等形式,我们排除了村规民约字符少于10个字的村庄。

②吕鹏,范晓光:《计算社会科学导论》,北京:清华大学出版社,2023年版,第135页。

终可用于生成村规民约文本类型分布数据的文本分类器。

文本分类器模型的训练、评估和运用过程如下。首先，根据村规民约中的常见分隔符（换行符号、分号和句号）将村规民约分为若干条款，设立每一个条款为最小样本单位。接着，将所有村庄的村规民约条款合并、打乱，随机从数据中挑选1000条条款进行人工标注并区分训练集、验证集和测试集。再次，使用常用优化算法和模型设定对BERT模型进行多轮分类任务训练，并分别在验证集和测试集上完成验证和测试。在模型分类准确率达标且模型在多轮测试中表现稳定后，保存为最终使用的文本分类器。最后，使用文本分类器对其余条款进行分类，并与村级数据整合，获得以“村庄（以及村庄各类属性）—村规民约—条款和标记”为基本结构的村规民约文本类型分布数据库。

通过Pytorch、Transformers构建基于BERT的文本分类器。训练上，本文采用不同的模型设置（包括批大小、学习率等）并在随机化的训练集和测试集上进行性能测试。本文最终选取了综合表现最好的模型，并随机抽取预测集结果进行人工审阅，基本满足区分不同类型文本的条件[①]。

为了更加直观展示本文模型分类的精准度，我们随机选择11条村规民约条款（见表1），可以发现，文本与预测标签的一致性较高。

表1　模型预测结果

文本	预测标签
每月一次的消防安全例行检查	灾害防治
多沟通、谈感情、远亲不如近乡邻	邻里关系
扔垃圾，要分类，无害减量环境好	环境卫生
兰干村所有村民均要执行村“两委”提出的维护社会稳定，发展庭院经济，“三区分离”、乡村振兴、稳岗就业、土地、果园流转、牛羊托管、人居环境整治、果园管理等工作决策部署	经济发展
集体主义，少数服从多数	治理与参与
建造房屋和临时建筑物，必须经村两委会和镇政府规划和审批	建筑与宅基地
移风易俗，不搞封建迷信，不搞礼金攀比，不搞高额彩礼，不搞跟风宴请，不强摊强派	风俗文明
热爱中国共产党，热爱祖国，热爱家园，遵纪守法、讲究文明、维护社会公德、维护社会治安	爱国守法
未经村委会批准，未取得“取土许可证"的不得在承包田上取土或卖沙	林耕农牧
举报、制止秸秆焚烧，在山野林地等随意取火（生火）行为，每次积2分	秸秆焚烧
优生优育，男女平等，尊老爱幼	家庭关系

三、分析发现

（一）研究发现

我们将每个村规民约的不同主题文本字数除以村规民约总字数得出每个主题在村规民约中权重，以此测量不同主题在村规民约中的重要程度（见图1）。通过对文本主题重要程度的分析，我们发现治理参与相关的条款在村规民约中是最常出现的，村规民约中平均有25%的文字是有关治理参与，例如“第三条积极为村庄建设捐献义务工”。爱国守法排在第二位，平均占比16%。邻里关系和风俗文明的条款大约占比10%，例如“为切实推行乡村治理移风易俗工作，根据村红白喜事理事会规定：本村红白喜事一切从简，席面费用不得超过300元，桌数不得超过5桌，人情往来不超过100元”，这表明在大操大办成为乡村治理的重要议题。建筑和宅基地也是村规民约的重要条款，主

①模型测试结果可联系作者获取。

要关于禁止村庄内部乱搭乱建、一户一宅、宅基地申请等相关的条款。家庭关系、经济发展、林耕农牧议题的平均权重约为5%左右。秸秆焚烧相关议题约占1%。

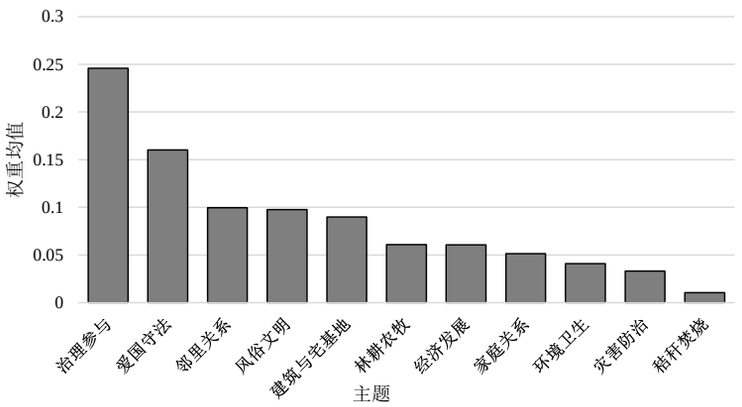


图1　不同主题平均权重

如图2所示,我们还对不同主题下村庄分布进行分析:超过75%的村庄村规民约中有关于爱国守法、风俗文明、邻里关系、治理参与的条款。爱国守法条款的普遍存在,反映了国家权威在基层社会的深度渗透和国家意识形态对乡村社会的有效灌输。国家权力通过制度和规范的方式在村庄层面上实现了有效的制度化延伸和文化渗透,村规民约成为乡村治理中国家合法性和权威性的载体。关于邻里关系的条款,体现了村庄对内部社会关系的重视和维系,这与乡土社会中强调熟人社会和人情网络的文化特质相符合。这种条款的广泛存在也反映了乡村社会对于维持内部社会秩序和防范社会冲突的迫切需求。治理参与条款在大多数村庄村规民约中的广泛出现以及治理参与的高比例显示出村民在乡村治理中的主体地位日益被重视,村规民约成为多数村庄推动村民参与公共事务的重要工具。

随着近几年在基层开展的环境整治工作,环境卫生条款也在村规民约中广泛出现,大约有66%村规民约包含环境卫生。61%村庄包含家庭关系,这表明通过村规民约对家庭关系进行规范,村庄希望促进家庭内部的和谐稳定,从而保障村庄社会的整体稳定性。

大约有63%的村规民约中包含建筑与宅基地相关条款。在农村地区,土地使用权和宅基地分配是村庄内部潜在冲突的主要来源之一。通过村规民约对建筑和宅基地进行规范,有助于减少土地利用中的矛盾和纠纷,提高土地资源的利用效率,维护村庄的社会秩序和经济发展稳定。此外,61%的村庄提及经济发展。58%的村庄提及林耕农牧,这也反映了村庄对于农牧业的规范和引导。虽然秸秆焚烧在村规民约中权重最低,但约有47%的村庄提及。

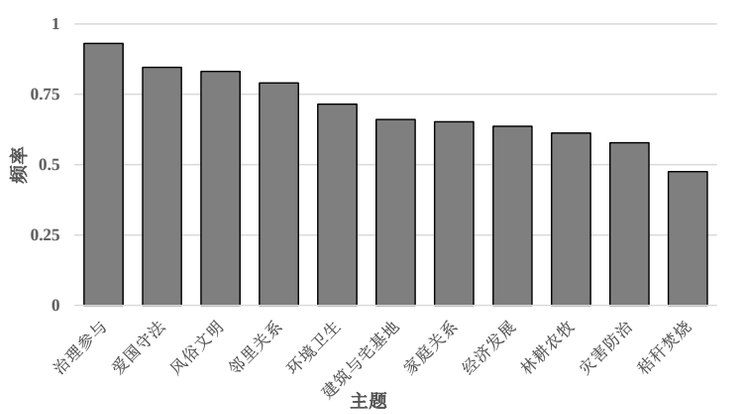


图2　不同主题村庄分布

（二）地区分析

我们将31个省直辖市、自治区划分为华东、华南、华中、华北、西南、东北、西北7个地区[①]，对主题权重进行分地区分析（见图3）。

治理参与和爱国守法作为所有地区村规民约的核心主题，体现了社会对集体规范和国家意识的普遍关注，这些主题在不同地区均被广泛重视，反映了社会对基本治理和法律遵循的共同期许。

然而，不同地区的具体主题分布显示了各地社会结构和经济活动的多样性。东北地区以邻里关系、林耕农牧和风俗文明为重要主题，在村规民约中，平均有15%的部分是关于邻里关系的表述，6%的则关于风俗文明与林耕农牧。这与该地区作为中国粮食主产区的农业基础密切相关。这里丰富的森林资源和耕地条件促使村规民约中更多地包含有关土地利用和乡村生活的条款。这种现象反映了社会结构中的农业主导性和对传统乡村习俗的依赖。

华北地区的邻里关系和风俗文明主题也较为突出，大约平均有14%的条款是关于邻里关系的表述，11%的条款关于风俗文明的条款。在华东、华南地区，风俗文明、邻里关系、建筑与宅基地则是关键条款。可能与这些地区经济发达、人口密集及城乡规划相关。

华中和西北地区在邻里关系和风俗文明方面的高权重表明，这些区域对社会秩序和文化传统的维护也具有重要性。特别是西北地区，风俗文明的权重接近于爱国守法，表明该地区在社会治理中同样高度重视文化习俗的融合。在西南地区，邻里关系和风俗文明的权重接近10%，显示出该地区对社会和谐与传统习俗的关注。这种权重分布的区域性差异，不仅揭示了地方社会结构的特点，也反映了不同地域在治理实践中的独特需求和文化价值。

总体而言，村规民约的主题权重分布不仅反映了地方经济基础和社会结构的差异，也揭示了各地区对社会治理和文化传承的不同重视程度。

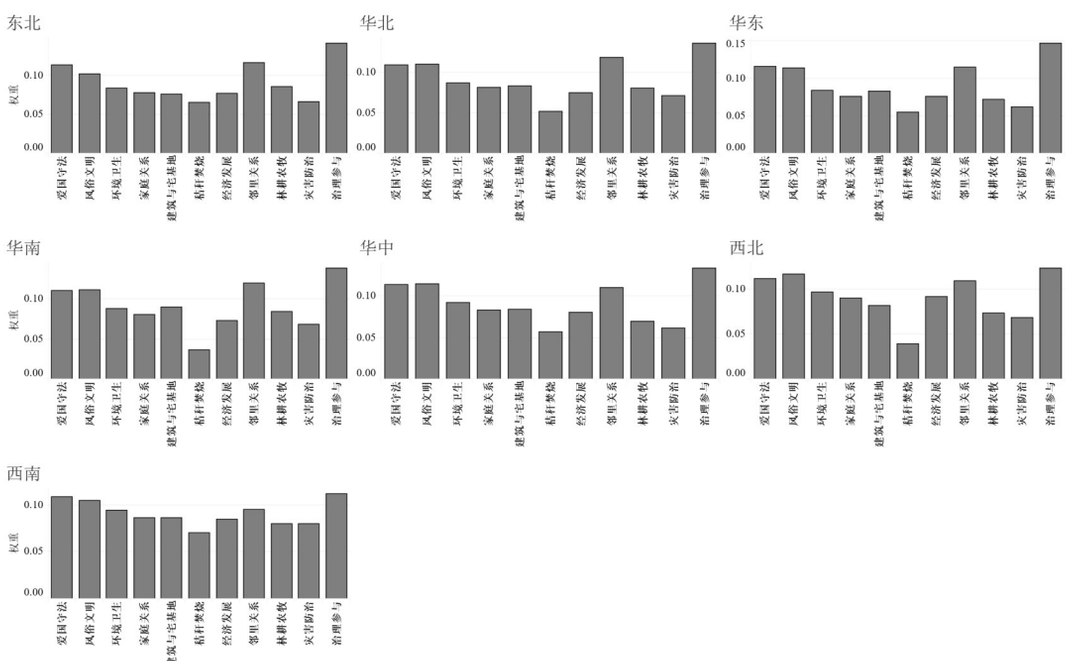


图3　不同区域主题权重分布

①华东地区：山东省、江苏省、安徽省、上海市、浙江省、江西省、福建省和台湾地区；华中地区：湖北省、湖南省、河南省；华南地区：广东省、广西壮族自治区、海南省、香港特别行政区、澳门特别行政区；华北地区：北京市、天津市、河北省、山西省和内蒙古自治区；西南地区：四川省、贵州省、云南省、西藏自治区、重庆市；东北地区：黑龙江省、吉林省、辽宁省；西北地区：陕西省、甘肃省、青海省、宁夏回族自治区、新疆维吾尔自治区。

如图4所示,通过对不同主题下村庄分布进行分析发现:治理参与在各个地区分布都是排名第一的,均高达80%以上,这反映出在中国不同地区的村庄中,治理参与具有普遍的重要性。秸秆焚烧都是比例最低的,其中华南地区,包含秸秆禁烧的村庄比例最少,仅有24%。在其他主题上不同地区的村规民约中主题的分布存在差异,这反映了各地区在社会文化、经济发展和治理实践上的多样性。

东北地区,有超过60%的村规民约中包含邻里关系、爱国守法、风俗文明。50%的村庄包含林耕农牧、环境卫生。大约47%的村庄村规民约条款中包含家庭关系、经济发展、建筑与宅基地,大约40%的村庄包含灾害防治和秸秆焚烧。

华北地区有超过75%的村庄村规民约中包含邻里关系条款,风俗文明与爱国守法约占69%,环境卫生主题占比55%,经济发展占比47%,建筑与宅基地约占52%,家庭关系、林耕农牧约占51%,经济发展与灾害防治约占45%,秸秆焚烧约占33%。

华南地区,邻里关系、风俗文明和爱国守法均超过70%,建筑与宅基地、环境卫生约占58%,林耕农牧占比56%,家庭关系占比53%,经济发展和灾害防治分别占比48%、45%。值得注意的是,华南地区在邻里关系和建筑与宅基地的条款相对比例明显高于其他区域,可能与当地社会结构和城市化进程有关。

华中地区,风俗文明、爱国守法、邻里关系占比超过75%。环境卫生占比62%,建筑与宅基地、家庭关系、经济发展分别占比57%、56%、54%。林耕农牧与灾害防治与秸秆焚烧分别占比47%、41%、38%。

西南地区,爱国守法、风俗文明占比都高达90%以上,邻里关系、环境卫生占比为81%,家庭关系、建筑宅基地和经济发展均占比72%左右,林耕农牧、灾害防治占比为68%。

西北地区,风俗文明、爱国守法占比超过80%以上,环境卫生占比70%,经济发展与家庭关系占比65%左右,建筑与宅基地占比59%,林耕农牧、灾害防治分别占比53%、50%。秸秆焚烧占比28%。

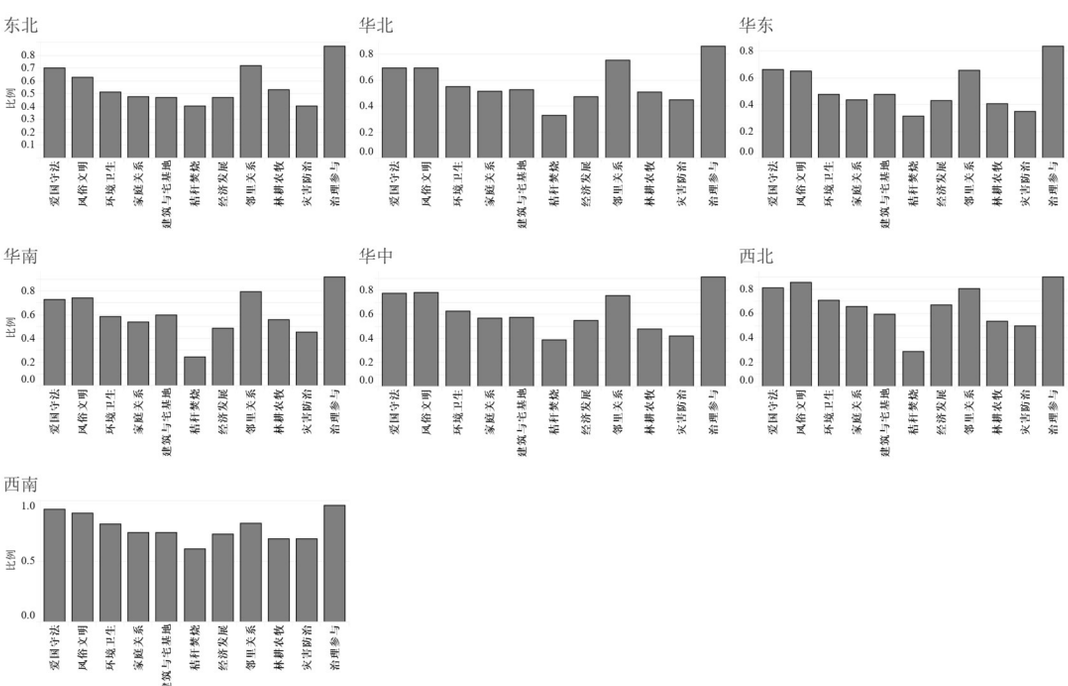


图4　不同主题村庄分布

在以上分析中，村规民约不同主题的分布体现了国家法律和政策的深度渗透，例如法律法规、治理参与的高频出现。这说明了村规民约受到国家意志的灌输。然而区域的差异性则表明了这种灌输并非单向的、全然主导的过程。地方社会的积极应对和选择性吸收，在吸纳国家意志的同时，也保持了自身的传统规范和地方风俗。这种双向互动关系体现了国家与社会关系理论中的“互动性”，即国家权力在基层的实施和执行过程中往往需要经过地方社会的过滤和再创造。

（三）基于村规民约的村庄类型划分

为了进一步了解不同区域村庄在村规民约上的异质性，我们采用LCA对村庄村规民约主题进行分类，通过不同分类拟合模型发现，4分类模型在不丧失过多的自由度的前提下，各项统计指标表现较好，同时，我们对4分类模型和5分类模型进行模差异性检验（P=1），这表明两个模型的对数似然值几乎相同，相较于4分类模型，5分类模型在拟合数据方面没有显著改善。因此我们将村庄分为4类（见表2）。

表2　LCA模型分类

模型	AIC	BIC	G2	X2	Df
2-Class	127359. 8	127534. 7	16283. 57	31502. 25	2024
3-Class	120583	120849. 2	9482. 834	24444. 84	2012
4-Class	115302. 2	115659. 7	4178. 033	6207. 174	2000
5-Class	114878. 3	115327. 1	3730. 153	5672. 181	1988
6-Class	113681. 3	114221. 3	2509. 103	4141. 785	1976
7-Class	113466. 8	114098. 1	2270. 657	3335. 394	1964

表3的数据显示，类型1的村庄在所有主题上都有非常高的概率，尤其在邻里关系、爱国守法、风俗文明、环境卫生等方面均达到1。这种村庄的治理模式全面而均衡，注重所有村规民约主题，体现出一种全方位的社区治理策略，我们将其命名为“完美和谐村”。类型2的村庄对“治理参与”表现出较高的重视（0. 71），而对其他主题的重视度相对较低。它们可能更关注村民的政治参与或村级事务管理，体现了一种专注于基层治理的模式，我们将其命名为“治理先锋村”。类型3的村庄在“爱国守法”和“风俗文明”上显示出极高的重视程度（接近1），而对其他主题（如建筑与宅基地、家庭关系等）则基本不重视，我们将其命名为“文明模范村”。类型4的村庄在多个主题上都有中等到高水平的重视，显示出一种均衡发展、多维治理的模式。它们既重视邻里关系、爱国守法、风俗文明等传统道德规范，也关注环境卫生、经济发展、家庭关系等实际事务，我们将其命名为“和谐共进村”。其中和谐共进型村庄在经济发展主题上弱于文明模范型村庄和完美和谐型村庄。

表3　不同类别在各主题的预测值

	邻里关系	爱国守法	风俗文明	环境卫生	经济发展	建筑与宅基地	家庭关系	林耕农牧	灾害防治	秸秆焚烧	治理参与
完美和谐村	1. 00	1. 00	1. 00	1. 00	0. 95	1. 00	1. 00	0. 99	0. 98	0. 90	1. 00
治理先锋村	0. 46	0. 37	0. 32	0. 09	0. 11	0. 15	0. 06	0. 17	0. 11	0. 06	0. 71
文明模范村	0. 02	1. 00	0. 99	0. 64	0. 99	0. 00	0. 01	0. 00	0. 00	0. 00	1. 00
和谐共进村	0. 95	0. 98	0. 97	0. 84	0. 50	0. 73	0. 78	0. 54	0. 48	0. 26	0. 99

进一步分析不同类型村庄在各地区的分布（见表4）。总体来看，全面治理型村庄在所有地区占比都是最少。在东北地区，村庄主要以和谐共进村为主。华东地区超过一半都是和谐共进村。在华中地区，不同类型村庄分布相对均衡，有32%的为治理先锋村，29%的文明模范村，34%的和谐共

进村。在华北地区,和谐共进村占比43%。华南地区和谐共进村庄占比37%,文明模范村庄占比30%。西北地区,治理先锋村庄占比最多,达到37%。西南地区治理先锋村庄比例最高,达到56%,和谐共进型村庄占比最少,仅有13%。

不同类型村庄在各地区的分布反映了中国农村社区治理模式的多样性和复杂性。东北、华东和华北地区的村庄更倾向于"和谐共进村"类型,这表明这些地区注重在邻里关系、爱国守法、风俗文明、环境卫生等多个维度上均衡发展,体现了一种全面、综合的村庄治理策略。这种模式可能与这些地区相对稳定的社会经济环境有关。相反,西南和西北地区的村庄更多表现为"治理先锋村",重视村民参与和基层治理事务,可能反映出这些地区在社会治理上面临更大的挑战,需要更高的治理参与度。而华中和华南地区的村庄类型分布相对均衡和多样化,既有强调治理参与的"治理先锋村",也有关注风俗文明的"文明模范村",以及注重多方面发展的"和谐共进村",这反映了这两个地区的社会经济文化环境的多样性。总体而言,这种村庄治理模式的区域差异,揭示了各地区在面对社会治理问题时的不同应对策略和社区组织结构的多元性。

表4　不同区域村庄类型分布

地区	完美和谐村	治理先锋村	文明模范村	和谐共进村
东北	3%	29%	21%	47%
华东	5%	24%	20%	51%
华中	5%	32%	29%	34%
华北	3%	32%	22%	43%
华南	6%	27%	30%	37%
西北	5%	37%	34%	25%
西南	8%	56%	23%	13%

四、结论

随着市场经济的不断发展,乡村社会的人际结构、家庭结构、生计模式、道德观念等方面发生了巨大的变迁[①]。理解中国乡村的各种正式制度,离不开理解其根源于社会的物质生活关系,去认识格式化制度与非格式化现实之间的张力。本文通过采用计算社会科学方法对近1.5万个村庄的村规民约进行文本挖掘,对村规民约背后的差异进行了初步的探索。大数据分析发现:首先,国家意志的融入成为村规民约的重要议题,无论是法律法规的主题,还是秸秆焚烧主题的高频出现,均反映出国家权力对基层社会的影响力在加强,村庄也在积极融合和吸纳国家法律。其次,在与国家趋同的同时,地方社会也在积极保持自身传统规范和灵活性。大量研究证明,国家的正式规则在基层应用往往会采取变通或者打折扣的方式运行[②]。村规民约中协调家庭关系、邻里关系、利益关系的内容成为不少村庄的重要主题,例如"竹子扩鞭到他人地界,则以土地所有为依据,属土地权属者所有。水田边,上12米内不能种树,下10米内不能种树。有荫山田可加6米"这些条款反映了村庄治理中地方性,而这种在社会实践中形成的各种地方性条款,是中国法治建设的重要本土性资源[③]。在中国广袤的农村社会中,村民不仅是政策的接受者,更是希望通过村规民约,努力参与到治理过程之中。这一过程不仅促进了法律的地方化,也强化了乡村社会的法治观念与秩序建构。最

①刘婷婷,罗强强:《简约抑或复合:乡村治理范式再思考》,《江淮论坛》,2023年第6期。

②向静林,邱泽奇,张翔:《风险分担规则何以不确定——地方金融治理的社会学分析》,《社会学研究》,2019年第3期。

③苏力:《法治及其本土资源》,北京:北京大学出版社,2015年版,第15页。

后，村庄内部风俗文明的强调，揭示了乡村社会在现代化转型过程中面临的文化冲突与自我调适。一方面，“人情”“大操大办”等问题的凸显，反映了传统习俗在现代价值观的映照下所展现的复杂面向；另一方面，将其纳入村规民约，则体现了乡村社会通过内部规范重构来应对文化变迁的努力。

区域分析的结果进一步丰富了我们对乡村治理多样性的理解。尽管国家拥有较高的社会渗透、动员能力，不同区域村规民约主题的差异性表明了国家意志在不同区域的渗透力和动员力是不同的，而代表国家意志“法制秩序”往往需要与乡土社会“礼治秩序”相结合[①]。其次，村规民约作为地方性知识的载体，在协调国家法律与乡土社会习惯法之间起着重要作用。不同区域的村规民约中主题的排序以及分布的差异都在一定程度上反映了“地方性”特征。地方治理主体在贯彻中央政策和法治的过程中，结合了本地实际情况和资源禀赋进行的创造性实践。最后，不同治理主题的差异在一定程度受到地方社会结构、经济发展水平、文化传统和历史背景的影响。本文从区域角度将各地的实践置于统一的背景下考虑，可以更加清楚的了解共通性和差异性在不同地区如何达到平衡[②]。这种差异性与共通性的共存，进一步丰富了我们对乡村治理多样性的理解。

本文在方法上结合了聚类分析和分类模型，创新地采用了预训练模型对大文本进行分析，有效规避了既往研究依赖单一无监督模型的不足，大幅提升了文本分析的精准度和可解释性。传统的词频分析方法通常只考虑词语的出现频率，难以捕捉词汇之间的复杂语义关系。而LDA主题模型虽然能够挖掘出文本的主题结构，但无法依照统一、稳定的标准对文本进行准确分析[③]。相比之下，预训练模型能够更好地理解上下文和词汇之间的深层关系，通过学习大量语料库中的语言特征，提供更准确和细致的文本分析结果。在数字社会的背景下，大数据和人工智能技术的发展为社会科学研究提供了新的方法和工具[④]。例如，预训练语言模型（如BERT、GPT等）能够通过多层神经网络结构，捕捉文本中词汇和句子之间的深层语义关系，从而在社会科学研究中获得更为细致和准确的分析结果。这种方法的应用使得研究者可以更加深入地理解复杂的社会现象，特别是在分析跨领域和多维度的数据时，具有显著的优势[⑤]。这种方法上的创新，推动了社会科学研究从传统的理论驱动分析向数据驱动的定量分析转变，为理解数字社会中的复杂现象提供了新的理论和实践工具[⑥]。

不可否认的是，许多村规民约仅仅是村庄治理中一种“形式”，它有时是应上级的要求而制定，甚至并未经过村民投票形成，在实践中并不总是能够得到严格的执行。但是，对于我们观察的基层组织而言，它们仍有不可多得的涵义。这是因为村规民约中的内容和不同主题总是会包含地方治理目标或者地方广泛认同的规则，这正是本文所关注的。村规民约是否得到村庄的严格执行，是我们未来需要观察的另外一个议题。在未来的研究中，我们也将村规民约的执行现状与村规民约的内容结合起来，以更加全面的了解村规民约在基层治理的现状和作用机制。

[责任编辑：王文娟]

①梁治平：《论法治与德治》，北京：九州出版社，2020年版，第186页。

②梁治平：《清代习惯法：社会与国家》，北京：中国政法大学出版社，1996年版，第184页。

③胡安宁：《以文本为基础的社会科学研究：从内容分析到算法模型》，《学术论坛》，2022年第1期。

④政光景，吕鹏：《生成式人工智能与哲学社会科学新范式的涌现》，《江海学刊》，2023年第4期。

⑤陈云松，吴晓刚，胡安宁等：《社会预测：基于机器学习的研究新范式》，《社会学研究》，2020年第3期。

⑥陈茁，陈云松：《计算扎根：定量研究的理论生产方法》，《社会学研究》，2023年第4期。