

# 我国人文社会科学学术图书被引影响因素研究\*

阮选敏, 吕冬晴, 张 培, 成 颖, 柯 青

**摘 要** 相较于期刊论文, 对图书被引影响因素的探究不足。文章以 CBKCI 收录的 1999-2009 年出版的图书为研究样本, 采用非参数检验及回归分析等统计分析方法, 探讨图书和作者的特征对图书被引的影响。研究发现: 书名长度、有无丛书项、所属学科、出版地和出版社等特征对图书被引有显著影响, 有无英文书名的影响不显著; 作者人数、机构类别、机构所在地、基金资助对图书被引存在显著影响; 10 个自变量对被解释变量回归模型的  $R^2$  徘徊于 0.2 左右, 显示更重要的影响因素有待于进一步发掘。

**关键词** 回归分析 CBKCI 作者特征 图书特征 人文社会科学 学术图书 被引影响因素

**引用本文格式** 阮选敏, 吕冬晴, 张培, 等. 我国人文社会科学学术图书被引影响因素研究[J]. 图书馆论坛, 2019 (5): 33-45.

## Factors Affecting Citations of Humanities and Social Sciences Books in China

RUAN Xuanmin, LYU Dongqing, ZHANG Pei, HENG Ying, KE Qing

**Abstract** Compared with journal articles, researches on factors affecting book citations are relatively few. So this needs further exploration. The paper makes use of books published from 1999 to 2009 and appeared in CBKCI as the research sample. Statistical analysis methods such as non-parametric test and regression analysis are used to explore whether publication related or author related features will influence book citations. Results show that (1) length of the title, part of a book series, subject area, the publisher and place of publication all have significant impact on book citations. But having an English title or not will not have any remarkable difference on citation. (2) The number of authors, their institutional affiliation, location of the institution and type of funding also significantly influence book citations. (3) By adjusting  $R^2$  of 10 selected variables in the regression model, the results are around 0.2. Therefore, further study is needed to identify the more important factors.

**Keywords** regression analysis; Chinese Book Citation Index; information on authors; information on publications

## 0 引言

学术文献被引频次作为辅助同行评议的重要手段<sup>[1]</sup>, 被广泛应用于科研评价工作中。但是, 学界对引用频次是否适合作为学术文献质量的评估指标尚未形成统一意见。据此, 探索学术文献

被引频次的影响因素, 理解引用的分布和意义就显得尤为重要。学术论文(以下简称论文)作为学术文献的重要类型, 其被引影响因素的研究已硕果累累。Tahamtan 等<sup>[2]</sup>对该主题进行了综述, 发现 28 个影响因素, 并将其分为三类: 论文、期刊、作者相关因素。在众多影响因素中, 哪些因

\* 本文系国家社会科学基金项目“施引者引用意向与文献计量视角的学术论文被引影响因素研究”(项目编号: 17BTQ014)研究成果。

素在论文被引中起到主要的作用以及如何利用其预测被引成为一个新的研究问题,负二项回归<sup>[3-4]</sup>和线性回归模型<sup>[5-7]</sup>是研究该问题的常用方法。

学术图书(以下简称“图书”)也是重要的学术文献类型。目前已经开展了多项有关图书被引的影响因素研究,研究者探讨了开放存取(Open Access)<sup>[8]</sup>、语言<sup>[8-9]</sup>、学科<sup>[8-10]</sup>、作者数量<sup>[11]</sup>、出版年份<sup>[9]</sup>、出版社<sup>[12]</sup>、是否为丛书<sup>[12]</sup>、是否为编著<sup>[9]</sup>等因素的影响。相较于论文,图书被引影响因素研究的广度和深度都有待进一步发掘。

相较于自然科学,图书对人文社会科学具有更重要的学术价值<sup>[13-18]</sup>,其被引在人文社科中的比例更高<sup>[15]</sup>。据此,本文开展中文人文社科图书被引影响因素的研究,具体内容包括:(1)书名长度、有无英文书名、有无丛书项、所属学科、出版地和出版社等6个图书相关因素对被引的影响;(2)作者数、机构类别、机构所在地和基金资助4个作者相关因素对图书被引的影响;(3)分析10个影响因素对图书被引的预测能力。

## 1 相关研究

目前对学术文献被引频次影响因素的研究以期刊论文为主,涉及论文、期刊、作者、参考文献等特征,而有关图书被引频次的研究较少。

### 1.1 作者特征

多项研究表明作者人数对学术文献的被引有显著影响。以论文为研究对象的大部分研究发现多作者论文能获得更多的引用<sup>[11][19-20]</sup>;不过,Bornmann等<sup>[21]</sup>的探究未发现二者间存在显著的相关关系,也有研究发现了二者间的负效应<sup>[22]</sup>。Thelwall等<sup>[11]</sup>选择BKCI 30个学科的图书为数据源,采用Mann Whitney检验,发现单作者和多作者图书的被引不存在显著差异。Abrizah等<sup>[9]</sup>发现在高被引图书中,单作者图书的比例显著高于多作者,从侧面说明多作者优势在图书中不明显甚至处于劣势。

作者的地区差异对被引存在影响。Sin<sup>[19]</sup>以6种图书情报学科的顶级期刊在1980-2008年发表的论文作为研究样本,逻辑回归结果显示作者为北美及北欧的论文被引高于东亚、东南亚、南

欧;作者来自中低收入国家的论文被引低于高收入国家。Peng等<sup>[23]</sup>将作者机构按照地理位置和文化相似度分成七个地区,发现来自美国的论文比来自中国大陆、日本等地区的被引频次要高,而与英国等地区没有显著差异。Willis等<sup>[24]</sup>发现来自美国的作者被引频次要高于亚洲地区的作者,欧洲在这方面的表现略逊于美国。

作者的机构属性也是影响被引的因素。作者机构的声望越高、大学的排名越靠前,论文的被引频次也越高<sup>[25]</sup>。Griffiths<sup>[26]</sup>针对团体作者的工作证实上述结论,其对联合国、欧盟等国际组织的研究发现,得益于良好的国际声望,联合国在国际研究领域的被引数量是其他国际组织的两倍。

基金论文具备引用优势得到了很多实证研究的支持。Amara等<sup>[27]</sup>对加拿大35个商学院学者的研究证实基金等级与被引和产出呈正相关关系。威尔鹏等<sup>[27]</sup>对Web of Science数据库收录的2010—2012年基础学科论文数据进行的实证研究表明,除逻辑学外,基础学科的基础论文具备引用优势。董建军<sup>[28]</sup>通过对《中国学术文献出版总库》中收录的各类基金论文和篇均被引频次的分析,发现基金论文整体上能够获得更高的被引量;不过,也有研究发现二者间的关系不显著<sup>[24]</sup>。

围绕作者数量的研究发现,针对论文与图书的结论存在明显差异,有必要探讨该变量对中文人文社会科学图书的影响;作者的地区差异、机构属性以及基金资助信息对于被引的影响目前均仅见到针对论文的结论,这三个因素对图书,尤其是中文人文社科类图书被引的影响有必要通过实证研究予以揭示。

### 1.2 文献特征

标题是文献的眼睛,是读者最先接触到的部分。不少研究分析了标题长度对论文被引的影响。一些研究发现,标题越长,被引频次越低<sup>[29-30]</sup>;但是在一些学科中,二者的关系正好相反。Van Wesel等<sup>[20]</sup>发现全科医学以及内科学(General & Internal Medicine)文献中,标题长度与论文被引呈现微弱的正相关。部分研究未发现二者的显性相关关系<sup>[31-32]</sup>。

学术文献的被引频次在不同学科间差异明

显。Tang<sup>[10]</sup>随机选择宗教学、历史学、心理学、社会学、数学和物理学 6 个学科 750 本图书,方差分析结果表明图书被引频次、被引率和 price 指数间有显著学科差异,即使是同一学科,论文与图书的被引差异也十分明显。Miettunen 等<sup>[33]</sup>发现精神病学四个子学科间的论文被引存在显著差异。

丛书和出版社是图书的独特特征,相关研究较少,仅有的研究表明<sup>[12]</sup>大学出版社出版的图书被引频次最高,而属于丛书的图书被引情况则因学科而异:工程技术和自然科学领域的丛书比非丛书图书的被引频次高;人文学科的结论正好相反,而社会科学未发现显著差异。

上述研究表明,同一影响因素对论文和图书被引的影响也存在一定差异;不论是从相关研究的数量来看,还是从已发掘的被引影响因素的数量来看,期刊论文的研究成果显著多于图书;图书作为学术交流的重要媒介,相较于论文,对其被引影响因素的研究明显不足。

### 1.3 回归分析

哪些特征是影响被引的主要因素,研究者常试图利用负二项回归和多元线性回归予以回答;部分研究还利用回归模型进行被引的预测研究。

引文数据符合泊松分布,但存在过度分散问题,负二项回归模型适合处理该类型的因变量。Bornmann 等<sup>[3]</sup>使用负二项回归模型,用期刊影响因子(JIF)、作者人数及评审分数等特征预测论文发表前十年的年被引频次和累计被引频次,发现 JIF 影响最大,而表示论文质量的同行评议分数对被引频次的影响较小。Didegah 等<sup>[4]</sup>利用负二项逻辑障碍模型(Negative Binomial-Logit Hurdle Model),以期刊和参考文献等特征为自变量预测论文的被引频次,发现 JIF 的影响依然最大。Snijder<sup>[6]</sup>使用负二项回归模型分析语种、学科和开放存取对图书被引的影响,发现三者均显著。

线性回归模型也是常用的模型之一。Yu 等<sup>[5]</sup>将论文作者、期刊及引用等 24 个特征作为自变量,线性回归发现被引文献数量、作者人数及期刊五年 IF 等能够预测论文的被引。Vanclay 等<sup>[6]</sup>将 JIF、期刊自引频次及论文类型等作为自变量,结果发现 JIF 仍然是最重要的影响因素,而作者

的 h 指数和作者人数等特征的影响低于预期。

回归分析通过控制其他变量以观察某一变量对因变量的影响,且多个变量对因变量的影响效果可以相互比较,因此,回归分析的结果往往让人信服。虽然负二项回归模型加符合引文的偏态分布特征,但也有学者指出使用负二项回归分析引文数据值得商榷<sup>[34]</sup>。故本文拟选择线性回归模型比较多个自变量对图书被引的影响。

## 2 数据与方法

### 2.1 数据集

本文选择 CBKCI 1999-2009 年出版的 2844 册图书为数据源。CBKCI 收录了 21 个人文社会科学的图书,覆盖范围广,收录的图书综合了被引频次以及专家意见,具有原创性以及较高的学术水平。选取 1999-2009 年出版的图书的原因是,本研究的因变量是图书的被引频次,发表时间较近的图书被引的可能性很小,把这些图书包含在分析数据中会削弱检验结果的信度<sup>[11]</sup>,选择 2009 年作为截止年可以为图书留足充裕的被引时间。

本研究没有选择 CBKCI 的被引数据,原因是该数据库的被引仅涉及为数不多的图书被引信息,没有图书的期刊被引数据,同时该数据库在短暂试用之后暂未继续开放,考虑到中文社会科学引文索引(CSSCI)的引文数据较为规范,是国内人文社科领域比较权威的数据库,同时引文数据的生产有一定时差,为确保数据的完整性,本文选择 1999-2014 年的引文数据作为图书被引的数据源。

### 2.2 数据清洗

为了获取每一本图书对应的被引信息,本文借鉴苏新宁<sup>[35]</sup>的数据处理方式,即用图书的题名、作者、出版年和出版社四个字段在 CSSCI 引文数据库中查找对应图书的被引信息。因为 CSSCI 数据库部分数据标引方式与 CBKCI 提供的图书元数据不完全一致,同时存在部分错误数据,所以分两个步骤确定图书的被引频次。第一步是精确匹配,即题名、作者、出版年和出版社四个字段能够在 CSSCI 中完全匹配,即认为是该图书的被引;第二步是模糊匹配,考虑到 CSSCI 中存在各种原因引起的数据错误,本文采取模糊匹配策

略：(1)题名：CSSCI 中被引文献的题名包含在 CBKCI 提供的图书题名中即匹配成功。(2)作者：CSSCI 中被引文献的作者包含在 CBKCI 提供的图书作者中即匹配成功。(3)年代：CSSCI 中被引文献的年代与 CBKCI 提供的图书出版年一致即匹配成功。(4)出版社：CSSCI 中被引文献的出版社包含在 CBKCI 提供的图书出版社中或者 CBKCI 提供的图书的出版社包含在 CSSCI 中被引文献的出版社中即也认为匹配成功。上述四个字段中有三个字段匹配成功，即纳入备选数据集。随之，对入选数据进行人工筛选，去除所选 2844 册图书之外的被引文献。经过上述两步操作之后，保留的被引记录有 111693 条，平均每本图书的被引频次为 39.27，年均被引频次为 3.66，零被引图书有 74 册。

### 2.3 自变量与因变量

CBKCI 对收录的图书进行标引，元数据包括书名、所属学科、作者名、作者机构、出版地、出版社、出版年、图书分类号、丛书项、ISBN、CIP、资助项目类型、图书简介以及作者简介等。考虑到 ISBN、CIP 以及作者名等对被引难觅潜在的影响，因此选择可能影响图书被引的 10 个特征作为自变量，其中作者机构还包括高等院校与大陆高校类型两个子类别(见表1)。

表 1 自变量信息

自变量名称	变量取值
书名长度	题名中的符号数量
作者人数	作者人数
学科	分 21 类：人文社科领域的 21 个学科(CBKCI)
作者机构类别	分 4 类：高等院校、政府机构、研究机构、其他
高等院校	分 4 类：“985 工程”院校、“211 工程”院校、普通院校、大陆地区以外院校
大陆高校类型	分 3 类：“985 工程”院校、“211 工程”院校、普通院校
作者机构所在地	分 4 类：美国、其他西方国家(英国、澳大利亚、比利时)、中国大陆、其他亚洲国家和地区(菲律宾、中国香港、中国澳门、中国台湾、日本、新加坡)
出版地	北京、上海、东北、华北、华东、华中、华南、西南、西北
出版社是否百佳	分 2 类：“全国百佳图书出版社”、非“全国百佳图书出版社”(是、否)
丛书	分 2 类：丛书、非丛书(是、否)
基金类别	分 5 类：无、其他基金、市厅级基金、省部级基金、国家级基金
英文名称	分 2 类：有、无

相关研究多以被引频次作为因变量，先前研究中出现过的指标有总被引频次、平均年被引频次以及不同年份的被引频次等。不同形式的被引频次指标仅仅考虑“量”，不能区分施引文献的“质”，因此，本文引入被引质量指标以实现从“质”的方面完善测度。参考 Thelwall 等<sup>[1]</sup>的做法，引入总被引频次作为因变量，类似地引入总被引质量指标；由于总被引频次未考虑时间的影响，引入图书的年均被引频次；为了消除半衰期对图书被引的影响，借鉴 IF5 的思想，再进一步引入五年被引频次和五年被引质量作为因变量(见表 2)。

表 2 因变量信息

因变量名称	变量取值
总被引频次	出版年至 2014 年间的总被引频次
年均被引频次	总被引频次/(2014-出版年)
总被引质量	公式(2)
五年被引频次	出版第二年至第六年间的被引频次
五年被引质量	公式(3)

首先，将 CSSCI 的来源期刊划分为三个等级：第一级是南京大学评选出的各人文社会学科的 31 种一流期刊；第二级是 1998-2018 年持续入选 CSSCI 的来源期刊；剩下的期刊归为第三级。依据施引文献发表的期刊将其划分为相应的三个等级，定义  $n_1$ ,  $n_2$ ,  $n_3$  分别为第一、二以及三级的施引文献量。其次，赋予不同等级的施引文献以不同的权重，参考李克特量表的反向编码方法，给一、二和三级施引文献分别赋予权重  $w_1=3$ ,  $w_2=2$ ,  $w_3=1$ 。最后，不同等级施引文献数量乘以对应的权重，求和之后即为被引质量得分，总被引质量( $Q_t$ )及 5 年被引质量( $Q_5$ )的计算见公式(2)和(3)。

$$\text{一年被引质量: } Q = \sum_{i=1}^3 w_i \cdot n_i \quad \text{公式(1)}$$

$$\text{总被引质量: } Q_t = \sum_{j=\text{出版年}}^{j=2014} Q \quad \text{公式(2)}$$

$$\text{五年被引质量: } Q_5 = \sum_{j=\text{出版年}+1}^{j=\text{出版年}+5} Q \quad \text{公式(3)}$$

## 2.4 数据分析方法

### 2.4.1 双变量分析

首先分别对自变量和 5 个因变量进行双变量分析。对于分类变量，采用非参数检验完成，即分别使用 Mann-Whitney U 和 Kruskal Wallis 检验两组及多组图书因变量间是否存在显著差异。

对于连续变量，使用 Spearman 相关系数判定自变量与各因变量间是否显著相关。

2.4.2 回归分析

(1)多元线性回归。表 1 中的变量为自变量，将类别变量转为虚拟变量(dummy variable)，回归策略采用进入(enter)；表 2 中的变量为因变量，考虑到引文数据的偏态分布，不适合直接应用于线性回归模型，参考 Vancly<sup>[6]</sup>的做法，对因变量进行对数变换，ln(原因变量 +1)。

(2)最优尺度回归。最优尺度变换可以解决统计建模时分类变量的量化问题。由于本文的自变

量以分类变量为主，因此，拟采用最优尺度回归以比较不同测量尺度的自变量对因变量的影响，即将经对数变换的因变量做离散处理，分别建立最优尺度回归模型。

3 结果

作者人数为 5 的图书仅 1 册，本文的检验方法多要求样本量在 5 以上，故去除该样本。

3.1 描述性统计与双变量分析

对于研究问题 1 和 2 的回答，本文采用描述性统计和双变量分析完成，结果见表 3 和表 4。

表 3 非参数检验与相关分析结果

自变量	因变量				
	总被引频次	年均被引频次	总被引质量	五年被引频次	五年被引质量
作者人数	-.044*	-0.034	-.044*	0.008	0.005
基金类别	117.929***	75.190***	118.713***	61.186***	69.869***
作者机构类别	12.750**	7.317	12.521**	11.590**	11.668**
高等院校	181.557***	195.455***	187.231***	175.878***	182.750***
大陆高校类型	157.590***	171.450***	161.350***	160.189***	165.039***
作者机构所在地	26.640***	29.352***	28.419***	17.949***	19.623***
书名长度	-0.102***	-0.073***	-0.097***	-0.038*	-0.038*
学科	352.605***	375.571***	341.532***	461.045***	445.085***
出版地	58.798***	80.949***	55.964***	97.944***	91.200***
英文名称	529291.500*	512701.500**	534382.500	501565.500***	509514.500**
丛书	889741.000***	902265.000***	885526.000***	902628.000***	896785.500***
出版社是否百佳	796705.000***	773050.500***	807830.500***	765023.000***	778678.000***

注：书名长度和作者人数对应单元格的值为 Pearson 相关系数，其余单元格为对应统计量；\*表示 p<0.05，\*\*表示 p<0.01，\*\*\*表示 p<0.001。

表 4 分类变量的均值表

变量		样本量	五年被引频次	变量		样本量	五年被引频次
学科	图书馆情报与文献学	60	32.75	出版地	北京	1858	19.84
	社会学	112	30.13		上海	345	19.62
	经济学	301	26.19		华南	79	15.52
	法学	415	25.74		华中	122	14.28
	体育学	27	22.81		华东	226	13.81
	政治学	252	20.33		华北	77	12.86
	管理学	202	18.53		东北	57	10.19
	人文经济地理	53	17.94		西南	55	8.93
	教育学	174	17.84		西北	24	6.63
	文化学	35	17.51	基金类别	无	1764	19.76
	外国语言文学	101	16.89		国家级	521	16.92
	心理学	20	16		其他	41	16.59
	新闻学与传播学	58	15.95		省部级	285	15.04
	哲学	253	15.76		市厅级	232	14.25
	考古学	44	11.48	作者机构所在地	其他西方国家	6	59.33
	历史学	195	10.18		美国	14	30.71
	中国语言文学	311	9.61		其他亚洲国家和地区	53	22.72
	马克思主义理论	26	9.19	高等院校	中国大陆	2770	18.04
	民族学	41	8.1		大陆地区以外高校	61	28.28
	宗教学	70	5.96		“985 工程”院校	1402	21.27
	艺术学	93	5.2		“211 工程”院校	377	13.2
作者机构类别	政府机构	75	21.91		普通院校	519	12.11
	其他	68	20.04	英文名称	无	2368	17.51
	高等院校	2359	18.15		有	475	22.06
	研究机构	341	17.98	丛书	否	1391	17.28
出版社是否百佳	否	1384	14.9		是	1452	19.22
	是	1459	21.47				

考虑到篇幅,表4仅列出了其他研究中使用较多的5年被引频次的均值。

### 3.1.1 作者相关因素

双变量分析的结果显示基金类别、高等院校、大陆高校类型以及作者机构所在地4个自变量与5个因变量间都呈显著的相关关系( $p<0.001$ );除了与年均被引频次的相关性未得到证实外,作者机构类别与其他四个因变量的相关关系显著( $p<0.01$ );作者人数与总被引频次和总被引质量为负相关( $p<0.05$ ),未发现其与年均被引频次、五年被引质量和五年被引频次三个因变量间存在显性的相关关系。

描述性统计发现:(1)无基金资助图书在所有因变量上的表现均高于有基金资助的图书;不同类别基金资助的图书,5个因变量均值的表现一致,由高至低均为其他基金>国家级基金>省部级基金>市厅级基金。(2)作者机构类别在5年被引频次上的表现依次为:政府机构>其他机构>高等院校>研究机构。如果样本仅考虑高校作者,则五个因变量由高至低均为非大陆地区高校>“985工程”高校>“211工程”高校>普通大陆高校。(3)第一作者的机构所在地的被引表现依次是其他西方国家>美国>其他亚洲国家和地区>中国大陆。

### 3.1.2 图书相关因素

双变量分析结果显示学科、出版地、丛书及出版社对5个因变量均存在显著影响( $p<0.001$ );英文名称在5个因变量上的表现存在较大差异,除对总被引质量的影响不显著外,其对另外4个因变量的影响均显著,显著性水平随5年被引频次、5年被引质量、年均被引频次、总被引频次顺

序递减;书名长度与5个因变量均呈微弱的负相关。

描述性统计发现:(1)图书被引的地区差异显著,根据因变量的均值,可将9个地区划分为3个等级,即北京、上海>华南、华中、华北、华东地区>西南、西北、东北地区。(2)在21个学科之中,图书馆情报与文献学等6个学科的5年被引大于20,且基本稳居前列;管理学等10个学科的5年被引频次介于10~20;中国语言学等5个学科始终处于后5位。(3)属于百佳出版社的优质出版社,其所出版的图书有突出的被引表现。(4)丛书中的图书可以获得更多的被引。(5)有英文名称的图书被引高于书名仅为中文的图书。

### 3.2 最优尺度回归分析

自变量中分类变量多达8个,进行最优尺度回归分析是一个合适的选择,为比较10个自变量对各因变量的影响程度,本文针对5个因变量分别构建最优尺度回归方程。5个方程均通过了显著性检验( $p=0.000$ ),拟合优度分别为:0.177、0.177、0.176、0.200、0.193,拟合效果较弱。各方程中所有变量的容差都在0.9以上,远大于0.1(见表5),说明变量之间的多重共线性低,最优回归方程结果准确。

在10个自变量中,作者人数仅对被引频次存在显著影响,对其余4个因变量影响不显著。有无英文名称对年均被引频次和五年被引频次的影响显著,对另外3个因变量未见显著影响。其余8个自变量对于5个因变量均存在显著影响(见表6)。定距变量:书名长度,以及两个定序变量:出版社是否百佳、作者人数,三者与因变量的系数均为负数,即书名长度越长、出版社非百佳、作者人数越多的图书,其在五年被引频次等

表5 总体样本最优尺度回归重要性与转换后容差

自变量	因变量									
	总被引频次		年均被引频次		总被引质量		五年被引频次		五年被引质量	
	重要性	转换后容差	重要性	转换后容差	重要性	转换后容差	重要性	转换后容差	重要性	转换后容差
基金类别	0.121	0.913	0.073	0.923	0.124	0.913	0.052	0.920	0.067	0.914
作者机构类别	0.010	0.974	0.003	0.972	0.013	0.973	0.009	0.973	0.011	0.971
作者机构所在地	0.042	0.987	0.054	0.988	0.047	0.987	0.028	0.988	0.035	0.988
出版地	0.065	0.984	0.078	0.966	0.061	0.986	0.082	0.957	0.086	0.950
学科	0.527	0.959	0.574	0.958	0.524	0.963	0.651	0.950	0.623	0.957
书名长度	0.061	0.975	0.035	0.974	0.056	0.976	0.011	0.970	0.017	0.970
作者人数	0.013	0.951	0.005	0.964	0.011	0.951	0.001	0.944	0.002	0.942
出版社是否百佳	0.123	0.956	0.145	0.942	0.112	0.961	0.129	0.933	0.117	0.928
丛书	0.033	0.972	0.022	0.975	0.047	0.973	0.026	0.973	0.034	0.973
英文名称	0.004	0.993	0.010	0.992	0.006	0.994	0.012	0.992	0.007	0.993

表 6 总体样本最优尺度回归标准系数

自变量	因变量				
	总被引频次	年均被引频次	总被引质量	五年被引频次	五年被引质量
基金类别	0.120***	0.089***	0.121***	0.081***	0.091***
作者机构类别	0.041***	0.025*	0.045***	0.042***	0.045***
作者机构所在地	0.083***	0.097***	0.087***	0.074***	0.081***
出版地	0.102***	0.104***	0.101***	0.111***	0.113***
学科	0.301***	0.314***	0.299***	0.357***	0.343***
书名长度	-0.102***	-0.081***	-0.096***	-0.054**	-0.063***
作者人数	-0.039**	-0.027	-0.032	-0.015	-0.016
出版社是否百佳	-0.129***	-0.141***	-0.124***	-0.135***	-0.127***
丛书	0.059***	0.047**	0.074***	0.059**	0.068***
英文名称	0.022	0.035*	0.027	0.039*	0.03

注：\*表示 p<0.05，\*\*表示 p<0.01，\*\*\*表示 p<0.001。

表 7 总体样本多元线性回归标准系数

自变量	因变量				
	总被引频次	年均被引频次	总被引质量	五年被引频次	五年被引质量
书名长度	-0.101***	-0.078***	-0.097***	-0.06***	-0.059***
出版社是否百佳	0.131***	0.144***	0.125***	0.138***	0.131***
丛书	0.061***	0.046**	0.067***	0.06***	0.066***
英文名称	0.026	0.042*	0.025	0.043*	0.039*
学科：哲学(基准组)					
学科：历史学	-0.076***	-0.084***	-0.065**	-0.079***	-0.074***
学科：马克思主义理论	-0.044*	-0.038*	-0.04*	-0.025	-0.027
学科：民族学	-0.037*	-0.038*	-0.036	-0.04*	-0.039*
学科：社会学	0.093***	0.117***	0.088***	0.099***	0.094***
学科：体育学	0.039*	0.039*	0.044*	0.06***	0.064***
学科：图书馆情报与文献学	0.067***	0.077***	0.074***	0.104***	0.105***
学科：艺术学	-0.15***	-0.14***	-0.157***	-0.16***	-0.169***
学科：政治学	0.046*	0.058*	0.044	0.07**	0.062**
学科：中国语言文学	-0.089***	-0.111***	-0.083***	-0.121***	-0.116***
学科：宗教学	-0.108***	-0.113***	-0.1***	-0.114***	-0.11***
学科：法学	0.102***	0.105***	0.098***	0.133***	0.124***
出版地：东北(基准组)					
出版地：华东	0.122***	0.112**	0.118**	0.1**	0.095**
出版地：华北	0.062*	0.06*	0.055*	0.045	0.039
出版地：华中	0.091**	0.088**	0.085**	0.074*	0.067*
出版地：华南	0.094***	0.087***	0.091***	0.081**	0.077**
出版地：西南	0.045	0.049*	0.042	0.047*	0.043
出版地：上海	0.221***	0.22***	0.212***	0.21***	0.202***
出版地：北京	0.228***	0.248***	0.217***	0.234***	0.219***
作者机构所在地：中国大陆(基准组)					
作者机构所在地：其他亚洲国家和地区	0.062***	0.071***	0.063***	0.06***	0.063***
作者机构所在地：其他西方国家	0.048**	0.056***	0.047**	0.045**	0.043*
作者机构所在地：美国	0.034*	0.04*	0.035*	0.023	0.023
作者机构类别：高校(基准组)					
作者机构类别：研究机构	0.035*	0.019	0.036*	0.024	0.025
基金类别：无(基准组)					
基金类别：市厅级	-0.053**	-0.044*	-0.049**	-0.032	-0.033
基金类别：省部级	-0.091***	-0.064***	-0.094***	-0.067***	-0.074***
基金类别：国家级	-0.092***	-0.061***	-0.09***	-0.059***	-0.061***
(常量)	***	***	***	***	***

注：\*表示 p<0.05，\*\*表示 p<0.01，\*\*\*表示 p<0.001；限于篇幅，表 7 未展示对 5 个因变量均无显著影响的自变量。

5 个因变量上的表现会越差。在 5 个方程中，学科、出版社是否百佳始终是重要性最高的 2 个自变量，而作者机构类别以及是否有英文名称的重

要性始终徘徊在 1%左右，居于重要性的后列。张文彤<sup>[36]</sup>建议，由于最优尺度回归主要给出的是变换后评分的分析结果，许多有用的信息被

隐含在变换过程中。使用者可以将最优尺度分析作为一种预分析手段,通过它快速发现各类别间的差异和联系,然后回到常规的建模方法,用合并相似类别、建立复杂的哑变量模型等方式得到更易于理解和应用的分析结果。根据该建议,本文为比较不同组别图书的被引情况,进一步构建包含虚拟变量的线性回归方程。

### 3.3 线性回归分析——总体

采用进入(enter)策略构建线性回归方程(见表7),5个回归方程的自变量与因变量间均存在显著的线性关系( $p=0.000$ )。线性回归方程和最优尺度回归方程的拟合优度基本相同。三个非定类变量:书名长度、出版社是否百佳的线性回归结果同最优尺度回归方程,作者人数对5个因变量均无显著影响。定类变量引入虚拟变量后发现,有多个自变量对5个因变量的影响均显著,相较于每个自变量的基准组:(1)学科层面,法学、社会学、体育以及图书馆情报与文献学的被引较高,艺术学、中国语言学、宗教学及历史学的被引较低;(2)出版于北京、上海、华东、华南、华中的图书被引较高,回归系数依次递减;(3)丛书中的图书、机构位于亚洲其他国家地区及其他西方国家作者编撰的图书被引较高;(4)受到省部级和国家级资助图书的被引较低。

除共性影响外,自变量对部分因变量也存在个性化的显著影响:(1)被引频次。调整后的 $R^2=0.176$ ,除上述共性的解释变量外(下同),回归模型中的马克思主义理论及民族学学科、受到市厅级资助的图书被引较低,而政治学学科、作者机构位于美国以及机构类型为研究机构的作者撰写的图书被引较高。(2)年均被引频次。调整后的 $R^2=0.178$ ,马克思主义理论和民族学两个学科及受到市厅级资助图书的效应同总被引频次;有英文名称、政治学学科、华北与西南地区出版以及作者机构位于美国的图书,被引较高。(3)总被引质量。调整后的 $R^2=0.17$ ,与前两个相比,该模型个性的自变量较少,马克思主义理论学科以及受到市厅级资助的图书的被引质量偏低,出版地位于华北、作者机构位于美国以及机构类型为研究机构的图书,被引质量较高。(4)五年被引频次。调整后的

$R^2=0.197$ ,民族学图书的五年被引较低,有英文名称、政治学学科以及西南地区出版的图书,五年被引较高。(5)五年被引质量。调整后的 $R^2=0.189$ ,民族学图书的五年被引质量较低,有英文名称及政治学学科的图书五年被引质量较高。

### 3.4 回归分析——高校

在研究样本中,作者来自高校的图书有2298本,占全部样本的80.83%,是数据集的主要组成部分,所以有必要探究不同类型高校对学术图书被引的可能影响。

按照约定俗成将高校类型具体化为“985工程”“211工程”以及普通高校。非参数检验的结果表明,作者来自不同类型高校的图书被引存在显著差异( $p=0.000$ )。总体而言,最优尺度回归分析结果发现,5个因变量的回归方程调整后的 $R^2$ 都在0.22以上,与全体样本相比有明显的提升,这说明加入了高校类型后,方程的整体解释度更高。具体地,与全体样本的最优尺度回归相比,高校的回归结果不同之处有:(1)有无英文名称对5个自变量的影响均不显著,说明先前有无英文名称对因变量的影响可能是由于没有控制高校类型而产生,并非由于该自变量自身的差异形成;(2)按照重要性程度来看,学科仍然是影响图书被引最重要的因素,高校类型取代出版社成为对图书被引影响的第二位因素,从而出版社是否百佳的影响效果被削弱。

为比较作者来自不同类型高校图书的被引差异,进一步采用包含虚拟变量的线性回归方程。分析发现,各方程的拟合优度与最优尺度回归方程的拟合优度基本相同,相较于总体样本的线性回归方程拟合优度有明显提升。在控制其他变量后,“985工程”院校与普通院校图书的被引差异十分明显,表现为:作者来自“985工程”高校图书的5个因变量均显著高于大陆普通高校;“211工程”高校与普通院校的图书被引差异也较为显著,具体而言,作者来自“211工程”高校图书的总被引质量和总被引频次显著高于大陆普通院校,但是二者在其他3个因变量上没有显著区别。综合上述分析可见,高校类型是影响图书被引的重要因素,“985工程”“211工程”院校



的学者所撰图书容易获得更多、更优质的引用。

## 4 讨论

### 4.1 图书相关因素

#### 4.1.1 书名长度

书名是一本图书的重要组成部分,体现了图书的主题,是读者判断是否阅读的最初依据之一。本文发现书名长度与被引频次和被引质量呈现微弱的负相关;Stremersch 等<sup>[29]</sup>以论文题名为分析对象的结果与本文的结论一致,得到类似结论的研究还有林佳瑜<sup>[30]</sup>等。可能的原因在于,论文或者图书的题名越长,题名中所包含的独立概念越多,则所研究的内容就越具体、研究范围越窄<sup>[37]</sup>,被引用的受众就越少。不过,有关题名长度对被引影响的结论并不一致,Rostami 等<sup>[31]</sup>、Jamali 等<sup>[32]</sup>的研究发现二者间相关关系不显著;Van Wesel 等<sup>[20]</sup>发现学科对该关系具有调节效应,比如在内科论文中,长题名对被引有利,而在社会学及应用物理学中,则短标题能够获得更多的被引。

#### 4.1.2 地区差异

不同出版地的学术专著被引情况存在显著差异,北京和上海地区出版的学术专著的被引频次和质量均较高,而西南、西北和东北的学术专著被引频次和质量偏低。一方面,地区间教育资源分配不均衡,比如西部地区的教育资源相对贫乏。根据 2017 年《中国统计年鉴》的数据,北京和上海 2016 年教育经费分别为 11171250 万元、10131153 万元;西北、西南和东北地区平均每省(市)的教育经费分别为 5395703.6 万元、9206546.8 万元和 7265483 万元,西北、西南和东北三地的教育经费明显少于其他地区。此外,在 39 所“985 工程”高校中,北京、上海、西北、西南和东北地区的“985 工程”高校数量分别为 8 所、4 所、4 所、3 所和 4 所,优质教育资源的悬殊差距直接引发地区间人才的不均衡。另一方面,北京地区拥有 53 家“百佳出版社”,上海地区有 7 家,与之相对的东北地区有 4 家,西南与西北地区仅陕西、四川两省有 4 家百

佳出版社,这说明北京与上海地区出版事业发展繁荣,而西部与东北地区出版事业的发展则相对缓慢。地区之间教育资源的不均衡和出版事业的发展差异是造成不同地区出版的图书影响力差异的重要原因。

#### 4.1.3 学科差异

徐贵水等<sup>[38]</sup>的研究发现,和生活越贴近、实用性越强、对人们的影响越大的学科,研究人数和学术成果也越多,可以吸引更多的被引;专业性与理论性强、离日常生活较远,研究人数和学术成果则会相对较少,进而得到被引也会相应减少。本研究证实了这一发现,不同学科图书的被引差异明显,兼具理论与应用性的学科,如法学、社会学、体育学、图书馆情报与文献学和政治学 5 个学科的图书被引领先于其他学科;长于理论的学科,如历史学、艺术学、中国语言学、宗教学、马克思主义理论等的图书五年被引频次等指标较低。本文的发现在论文的被引影响因素研究中也得到了证实,比如,Tang<sup>[10]</sup>以及 Miettunen 等<sup>[33]</sup>的研究发现,不同学科间的引用存在显著差异,Bornmann 等<sup>[21]</sup>发现研究面狭窄的学科获得被引的概率相对较低;本文也发现,历史学大类中的考古学,作为一个研究较为专深的学科代表,其被引在所有学科中处于后列。

#### 4.1.4 出版社质量

国内未见正规的出版社排名,所以本文仅采用是否为百佳出版社这一指标区分出版社的质量。由分析结果可知,百佳出版社出版的图书在被引频次和被引质量上均显著高于非百佳出版社,说明优秀出版社和高品质图书有强相关性。核心出版社具有信息密度高、文献新颖、文献利用率高和学术声誉高等良好特征<sup>[39]</sup>。Torres-Salinas 等<sup>[12]</sup>的研究表明大学类出版社出版图书的被引频次比其他出版社高,他将原因归结于大学出版社的声望高。本研究中的百佳出版社在国内出版行业有着良好口碑,即证实了 Torres-Salinas 等<sup>[12]</sup>的研究结论。人文社科学者倾向于在该领域声望最高的出版社出版自己的作品<sup>[9]</sup>,从而优质图书与高声望出版社的有机结合实现了双赢,带

动了源于高质量出版社图书的高被引。反过来,实践中出版社的声誉也用于图书质量的评估,比如作为图书馆员挑选图书的标准之一<sup>[40]</sup>。

#### 4.1.5 是否是丛书

Torres-Salinas 等<sup>[12]</sup>的研究表明人文艺术学科、非丛书图书的被引频次更高;社会学科,图书是否具有丛书标识对被引频次未发现影响;工程技术与自然科学的结论与人文艺术正好相反。本文的研究显示,丛书类图书的被引高于非丛书类图书。两项研究结果的差异可能源自引文数据的不同,Torres-Salinas 等<sup>[12]</sup>研究使用的引文数据来自图书引文(BKCI)和期刊引文(WOS)两部分,本文的数据为 CSSCI 的来源期刊对图书的引用,不包括图书之间的引用关系,因此得到的结果可能有所偏差。相较于非丛书,丛书内容具有更好的系统性和完整性,与独立著作相比,其丰富的信息量更容易得到读者的认可进而获得更多的引用。

#### 4.1.6 英文名称

非参数检验结果表明,有英文名称的图书的被引高于无英文名称的图书。但在回归方程中,控制了其他变量之后,有无英文名称仅对总被引频次和总被引质量有显著影响,不过系数很小,在高校样本的最优尺度回归结果中,有无英文名称对 5 个因变量的影响均不显著,表明有无英文名称对被引的影响可以忽略。

### 4.2 作者相关因素

#### 4.2.1 基金资助

针对不同资助级别对被引的影响,Amara 等<sup>[25]</sup>对加拿大 35 个商学院学者的研究发现基金等级与被引和产出呈正相关关系,本文的研究也得到了类似的结论,即受到市厅级、省部级、国家级基金资助的图书,其被引呈现递增趋势。对于有无基金资助对被引的影响,董建军<sup>[28]</sup>发现基金论文从整体上能够获得更高的被引;赵星等<sup>[41]</sup>的研究也证实了基金资助对被引的正向影响;本文未能证实董建军<sup>[28]</sup>和赵星等<sup>[41]</sup>在以论文为样本的研究中发现的现象;本文发现,无基金资助的图书能够获得更多的被引,该发现出乎笔者的预料。

对论文来说,基金项目筛选过程严格<sup>[27]</sup>,资

助的经费能为研究提供外部支持,有助于提高论文的质量<sup>[42]</sup>,所以基金资助论文容易获得更多的引用,可用于解释赵星等<sup>[41]</sup>的研究结论。本文如果采集所有的人文社科图书,那么基金资助图书的被引表现优于非资助图书值得期待;不过,本文数据集来自 CBKCI,来源图书源于专家推荐及 CSSCI 筛选的高被引图书,也就是说,总体而言本文数据集中的图书均为高质量,那么要思考的问题是:在高质量图书中,非基金资助图书的质量是否会高于基金资助图书?人文社科研究,尤其是人文科学与自然科学相比,对基金资助并没有特别的依赖,高质量图书的产出更多地依赖于学者对研究主题长期的浸淫,而基金资助图书有明确的研究时间要求,“十年磨一剑”难以体现。

#### 4.2.2 作者人数

Thelwall 等<sup>[11]</sup>发现在 12 个学科中,单作者的图书被引频次高于多作者;Abrizah 等<sup>[9]</sup>研究证实高被引图书中,单作者图书比例显著高于多作者;本研究也发现了这一现象,即作者人数越多,图书的总被引频次和被引质量越低。以论文为研究对象的工作中,发现多作者论文的被引往往高于单作者<sup>[43-45]</sup>。Van Dalen 等<sup>[46]</sup>认为多作者合作能够互相补充、取长补短,合作带来的专业化和分工优势是高引用的合理解释。作者数对不同文献类型被引影响的差异,意味着合作的优势在短出版物中更加明显,而随着出版物篇幅的增加,到图书这一类型的出版物时,合作的劣势已经不复存在<sup>[11]</sup>,本研究甚至还发现了合作“劣势”。

多作者合作对图书与论文被引影响的差异可能源于合作方式的不同。论文的篇幅通常较短,合作更多地体现在研究过程中,能够获得 Van Dalen 等<sup>[46]</sup>所说“专业化和分工优势”,而论文初稿的写作通常由一人完成,其他研究者更多地参与修改,整个过程中不同作者互相取长补短,收“1+1>2”之效,进而显著提升科研成果的质量。图书的篇幅远大于论文,科研合作的粒度显著增大,在研究实施阶段的科研分工也不同于细粒度的单篇论文,多数采用粗粒度模块化的分工方式,在成文过程中,也基本按照章节的划分由每

个研究者独立完成,从而相对于论文,作为科研成果的图书存在科研合作与写作合作两个阶段,由于每位研究者研究能力与写作能力的差异,整体研究质量与图书质量的保证有赖于项目负责人扎实、细致的工作,而这显然有难度。经济学中的规模经济<sup>[47]</sup>理论可用于解释论文中多作者合作带来的被引提高,而规模不经济理论则可用于解释图书的多作者合作引发的被引降低现象。

#### 4.2.3 作者机构所在地

本研究发现作者来自美国等西方国家的图书被引显著高于大陆地区,该结论与现有的以论文为对象的研究结果吻合。**Leimu**等<sup>[48]</sup>发现,相较于母语为非英语的国家,母语为英语国家的作者,其论文被引更高。**Sin**<sup>[19]</sup>证实作者位于北美和北欧地区的论文被引高于东亚等地区。**Peng**等<sup>[23]</sup>发现作者来自美国和英国的论文比来自中国大陆、日本等地区的被引频次高。**Willis**等<sup>[24]</sup>也发现美国作者的被引频次要高于亚洲地区。可见学术文献的被引情况在地区分布上存在“马太效应”,现阶段欧美等西方国家拥有的优秀学者和学术资源领先于其他国家与地区,优秀学者集中度高的国家会产出更多优秀的科研成果,吸引了更多的被引<sup>[19]</sup>。此外,收入水平等经济因素也是造成被引地区差异的原因。**Sin**<sup>[19]</sup>证实了作者来自中、低收入国家的论文被引要低于高收入国家;**Tahamtan**等<sup>[2]</sup>认为不同国家的科研水平、基金支持力度不同,造成了论文质量的差异,从而引起被引的不同。

#### 4.2.4 研究机构类别

本研究发现,作者来自中国大陆以外地区的高校,其图书被引显著高于大陆地区,原因已在“作者机构所在地”一节阐明。在大陆高校中,高校类型对于图书被引的影响十分显著,作者来自“985工程”“211工程”及普通院校的图书被引依次降低,即作者所属的院校排名越靠前,图书被引越高,这与多项研究不谋而合,例如**Amara**等<sup>[25]</sup>的研究。该结果合理的解释是,与排名较低的大学相比,排名靠前的大学会提供更有利于培养学术生产力的物质、智力和社会激励措施等资源<sup>[25]</sup>,从而吸引了更多的优秀学者加盟,出现更多被引的优质成果则是水到渠成之事。

### 4.3 回归分析

对样本总体和高校样本的最优尺度回归发现,学科、高校类型、出版社是否百佳是影响图书被引最重要的3个因素。据此,学科对被引的影响应引起科研管理部门的重视,在科研评价工作中,应当充分考虑绝对被引频次受到学科等因素的影响,在学科间不能采用绝对值进行比较<sup>[49]</sup>。马太效应提示人们不仅仅引用对自己有帮助的论文,还倾向于引用著名学者的成果<sup>[9]</sup>,与这种效应类似,本研究发现了学者不仅会引用与自身研究相关的图书,还倾向于引用声誉高的出版社、排名靠前高校产出的图书。

最优尺度回归和线性回归结果显示,本文的10个自变量对图书被引的拟合优度较低,说明本研究涉及的自变量尚不是影响图书被引的主要因素。以论文为数据源的研究证实论文质量是影响其被引的主要因素。**Yu**等<sup>[9]</sup>采用论文前两年的被引频次等引用特征作为衡量论文质量的指标,回归分析结果证实了论文质量是影响被引最重要的因素,在去除论文质量的引用相关指标后,模型的拟合优度由0.674迅速下降至0.177,说明论文质量对被引的重要性。目前一些研究表明图书的内在质量和形式质量是衡量图书质量的重要指标<sup>[50-52]</sup>。相较于论文,图书尚缺少类似于JIF等约定俗成的用于评价论文内在质量的相关指标。

## 5 结语

本文以CBKCI收录的1999-2009年出版的图书为研究样本,探讨图书被引的影响因素。非参数检验和Spearman相关分析表明,对五个因变量均有显著影响的特征有:学科、作者机构所在地、出版地、是否为丛书、出版社是否百佳、基金类别、高等院校、大陆高校类型和书名长度;作者数仅对图书的总被引频次和总被引质量有显著影响;有无英文名称对除总被引质量外的其余四个因变量有显著影响;除了年均被引频次之外,作者机构类别对其余的四个因变量均有显著的影响。结合回归分析的结果可以认为有无英文名称对图书被引的影响可以忽略。最优尺度回归和线性回归结果说明学科、出版社质量和高校

类型是影响图书被引最重要的因素。回归方程显示自变量对被解释变量的拟合优度较低, 后继工作拟进一步探究图书质量、作者特征等因素对图书被引的影响。本研究的不足之处是, 选取的数据集源于CBKCI, 属于高质量图书, 没有涵盖人文社科图书; 得到的基金资助图书被引低于非基金资助图书的结论有待于在全集中继续检验; 引文数据来源于CSSCI, 缺乏图书的引文数据, 待CBKCI开放后进行补充研究。

## 参考文献

- [1] Kousha K, Thelwall M, Rezaie S. Assessing the Citation Impact of Books: The Role of Google Books, Google Scholar, and Scopus[J]. Journal of the American Society for Information Science and Technology, 2011, 62 (11): 2147-2164.
- [2] Tahamtan I, Afshar A S, Ahamdzadeh K. Factors Affecting Number of Citations: A Comprehensive Review of the Literature[J]. Scientometrics, 2016, 107 (3): 1195-1225
- [3] Bornmann L, Leydesdorff L. Does Quality and Content Matter for Citedness? A Comparison with Para-Textual Factors and Over Time [J]. Journal of Informetrics, 2015, 9 (3): 419-429.
- [4] Didegah F, Thelwall M. Which Factors Help Authors Produce the Highest Impact Research? Collaboration, Journal and Document Properties [J]. Journal of Informetrics, 2013, 7 (4): 861-873.
- [5] Yu T, Yu G, Li P Y, et al. Citation Impact Prediction for Scientific Papers Using Stepwise Regression Analysis [J]. Scientometrics, 2014, 101 (2): 1233-1252.
- [6] Vanclay J K. Factors Affecting Citation Rates in Environmental Science [J]. Journal of Informetrics, 2013, 7 (2): 265-271.
- [7] Royle P, Kandala N-B, Barnard K, et al. Bibliometrics of Systematic Reviews: Analysis of Citation Rates and Journal Impact Factors[J]. Systematic reviews, 2013, 2: 74.
- [8] Snijder R. Revisiting an Open Access Monograph Experiment: Measuring Citations and Tweets 5 Years Later[J]. Scientometrics, 2016, 109 (3): 1855-1875.
- [9] Abrizah A, Thelwall M. Can the Impact of Non-Western Academic Books be Measured? An Investigation of Google Books and Google Scholar for Malaysia[J]. Journal of the Association for Information Science and Technology, 2014, 65 (12): 2498-2508.
- [10] Tang R. Citation Characteristics and Intellectual Acceptance of Scholarly Monographs [J]. College & Research Libraries, 2008, 69 (4): 356-369.
- [11] Thelwall M, Sud P. No Citation Advantage for Monograph-Based Collaborations? [J]. Journal of Informetrics, 2014, 8 (1): 276-283.
- [12] Torres-Salinas D, Robinson-Garcia N, Cabezas-Clavijo A, et al. Analyzing the Citation Characteristics of Books: Edited Books, Book Series and Publisher Types in the Book Citation Index[J]. Scientometrics, 2014, 98 (3): 2113-2127.
- [13] Zuccala A, Van Leeuwen T. Book Reviews in Humanities Research Evaluations [J]. Journal of the American Society for Information Science and Technology, 2011, 62 (10): 1979-1991.
- [14] Williams P, Stevenson L, Nicholas D, et al. The Role and Future of the Monograph in Arts and Humanities Research[J]. Aslib Proceedings, 2009, 61 (1): 67-82.
- [15] Glanzel W, Schoepflin U. A Bibliometric Study of Reference Literature in the Sciences and Social Sciences [J]. Information Processing & Management, 1999, 35 (1): 31-44.
- [16] 叶继元. 改革开放 30 年学术发展的主要特点和重要成果探视——基于图书被引用的分析[J]. 云梦学刊, 2008 (4): 24-27.
- [17] Thompson J W. The Death of the Scholarly Monograph in the Humanities? Citation Patterns in Literary Scholarship[J]. Libri, 2002, 52 (3): 121-136.
- [18] Verleysen F T, Ossenblok T L B. Profiles of Monograph Authors in the Social Sciences and Humanities: An Analysis of Productivity, Career Stage, Co-Authorship, Disciplinary Affiliation and Gender, Based on a Regional Bibliographic Database [J]. Scientometrics, 2017, 111 (3): 1673-1686.
- [19] Sin S C J. International Coauthorship and Citation Impact: A Bibliometric Study of Six LIS Journals, 1980-2008 [J]. Journal of the American Society for Information Science and Technology, 2011, 62 (9): 1770-1783.
- [20] Van Wesel M, Wyatt S, Ten Haaf J. What a Difference a Colon Makes: How Superficial Factors Influence Subsequent Citation[J]. Scientometrics, 2014, 98 (3): 1601-1615.
- [21] Bornmann L, Schier H, Marx W, et al. What Factors Determine Citation Counts of Publications in Chemistry Besides Their Quality? [J]. Journal of Informetrics, 2012, 6 (1): 11-18.

- [22] So M, Kim J, Choi S, et al. Factors Affecting Citation Networks in Science and Technology: Focused on Non-Quality Factors[J]. *Quality & Quantity*, 2015, 49 (4): 1513-1530.
- [23] Peng T Q, Zhu J J H. Where You Publish Matters Most: A Multilevel Analysis of Factors Affecting Citations of Internet Studies[J]. *Journal of the American Society for Information Science and Technology*, 2012, 63 (9): 1789-1803.
- [24] Willis D L, Bahler C D, Neuberger M M, et al. Predictors of Citations in the Urological Literature[J]. *Bju International*, 2011, 107 (12): 1876-1880.
- [25] Amara N, Landry R, Halilem N. What Can University Administrators Do to Increase the Publication and Citation Scores of their Faculty Members? [J]. *Scientometrics*, 2015, 103 (2): 489-530.
- [26] Griffiths D N. The Citation of United Nations Information Resources in International Relations Literature [J]. *Behavioral & Social Sciences Librarian*, 2008, 26 (4): 1-19.
- [27] 威尔鹏, 叶鹰. 基础学科论文的基金资助引用优势研究[J]. *大学图书馆学报*, 2015 (6): 11-16.
- [28] 董建军. 中国知网收录的基金论文资助现状和被引情况分析[J]. *中国科技期刊研究*, 2013 (2): 307-312.
- [29] Stremersch S, Camacho N, Vanneste S, et al. Unraveling Scientific Impact: Citation Types in Marketing Journals[J]. *International Journal of Research in Marketing*, 2015, 32 (1): 64-77.
- [30] 林佳瑜. 论文标题与下载和引用的关系[J]. *大学图书馆学报*, 2012 (4): 14-17.
- [31] Rostami F, Mohammadpoorasl A, Hajizadeh M. The Effect of Characteristics of Title on Citation Rates of Articles[J]. *Scientometrics*, 2014, 98 (3): 2007-2010.
- [32] Jamali H R, Nikzad M. Article Title Type and Its Relation with the Number of Downloads and Citations [J]. *Scientometrics*, 2011, 88 (2): 653-661.
- [33] Miettunen J, Nieminen P. The Effect of Statistical Methods and Study Reporting Characteristics on the Number of Citations: A Study of Four General Psychiatric Journals[J]. *Scientometrics*, 2003, 57 (3): 377-388.
- [34] Thelwall M, Wilson P. Regression for Citation Data: An Evaluation of Different Methods [J]. *Journal of Informetrics*, 2014, 8 (4): 963-971.
- [35] 苏新宁. 我国人文社会科学图书被引概况分析——基于 CSSCI 数据库[J]. *东岳论丛*, 2009 (7): 5-13.
- [36] 张文彤. SPSS 统计分析高级教程[M]. 北京: 高等教育出版社, 2004.
- [37] 肖学斌, 柴艳菊. 论文的相关参数与被引频次的关系研究[J]. *现代图书情报技术*, 2016 (6): 46-53.
- [38] 徐贵水, 孙莹莹. 我国百年学术发展特点初探: 基于读秀中文图书被引用报告的分析[J]. *情报杂志*, 2011 (11): 47-51.
- [39] 蒋志强, 徐文贤. 关于“核心出版社”的理论探索[J]. *图书情报工作*, 2002 (4): 62-65.
- [40] Jordy M L, Mcgrath E L, Rutledge J B. Book Reviews as a Tool for Assessing Publisher Reputation[J]. *College & Research Libraries*, 1999, 60 (2): 132-142.
- [41] Zhao S X, Lou W, Tan A M, et al. Do Funded Papers Attract More Usage? [J]. *Scientometrics*, 2018, 115 (1): 153-168.
- [42] 朱大明. 基金论文与非基金论文作者自引对比分析[J]. *编辑学报*, 2009 (3): 274-275.
- [43] 王海涛, 谭宗颖, 陈挺. 论文被引频次影响因素研究——兼论被引频次评估科研质量的合理性[J]. *科学学研究*, 2016 (2): 171-177.
- [44] 孟凡蓉, 郑烨, 杨若愚. 科技管理核心期刊论文被引频次影响因素——基于国内 5 本高被引期刊的实证研究[J]. *出版科学*, 2016 (2): 57-63.
- [45] Gazni A, Didegah F. Investigating Different Types of Research Collaboration and Citation Impact: A Case Study of Harvard University's Publications [J]. *Scientometrics*, 2011, 87 (2): 251-265.
- [46] Van Dalen H P, Henkens K. What Makes a Scientific Article Influential? The Case of Demographers [J]. *Scientometrics*, 2001, 50 (3): 455-482.
- [47] 顾江. 规模经济论[M]. 北京: 中国农业出版社, 2001.
- [48] Leimu R, Koricheva J. What Determines the Citation Frequency of Ecological Papers?[J]. *Trends in Ecology & Evolution*, 2005, 20 (1): 28-32.
- [49] 周晓英, 董伟, 朱小梅, 等. 图书馆情报学高影响力论文特征及所反映的学科差异分析[J]. *中国图书馆学报*, 2012 (4): 99-109.
- [50] 牛太臣, 曾勉之. 全面提高图书质量管理的尝试与探索[J]. *出版发行研究*, 1993 (1): 37-40.
- [51] 王大达. 论图书质量的量化评估和调控[J]. *出版发行研究*, 1993 (3): 30-37.
- [52] 代根兴, 周晓燕. 出版社与图书馆图书评价标准比较研究[J]. *大学图书馆学报*, 2013 (5): 36-40.
- 作者简介** 阮选敏, 吕冬晴, 张培, 南京大学信息管理学院硕士研究生; 成颖 (通信作者, chengy@nju.edu.cn), 柯青, 南京大学信息管理学院教授、博士生导师。
- 收稿日期** 2018-09-13
- (责任编辑: 刘洪; 英文编辑: 杨继贤)