

企业区位视角下的数据误差研究 ——以A股上市公司为例

胡国建^{1,2}, 陆玉麒³, 钟业喜^{1,2*}

(1. 江西师范大学地理与环境学院, 南昌 330022; 2. 江西师范大学鄱阳湖湿地与流域研究教育部重点实验室, 南昌 330022;
3. 南京师范大学地理科学学院, 南京 210023)

摘要:减少统计数据误差, 提高数据质量对保障政府和企业的合理决策以及大量实证研究结论的可靠性具有重要意义。论文鉴于众多社会经济数据按企业注册地址进行统计, 提出造成数据统计误差的新视角——企业实体位置与注册地址的空间分离, 并以A股上市公司为例, 考察公司数量、公司指标、发展环境指标(地区指标)以及地理距离等4类数据在省级尺度上按注册地址统计的误差幅度。研究显示, 若按注册地址统计: 各省上市公司数量的误差率平均为14.23%, 6项公司指标归并的误差率平均为24.84%, 且不同省份、不同指标的误差率差异较大; 各行业上市公司按注册地址统计其8项地区指标平均误差率为3.04%, 误差率较小; 许多公司在总部所在地聘用会计师事务所, 造成按注册地址计算上市公司到会计师事务所的地理距离会明显高于实际值。总体而言, 企业实体与注册地址的分离现象对各类按后者统计的社会经济数据造成了不可忽视的误差影响, 经济地理学对此应予以重视。

关键词:企业实体; 注册地址; 总部; 上市公司; 数据误差

企业的地理分布对社会经济发展有着举足轻重的作用, 也是造成人口流动、产业转移、贫富差距、环境污染等诸多社会现象的重要原因^[1]。因此, 企业相关的社会经济数据在科学研究和政策制定中具有重要地位, 而企业位置的明确是统计企业相关数据的前提。目前, 许多社会经济数据中的企业位置标准是注册地址, 即企业向工商部门申请成立时登记并记载于营业执照中的位置。如Wind数据库和Choice数据库依据注册地址划分上市公司所属的省、市、县(区), 并以此统计各省市的公司数量; 北京大学企业大数据研究中心发布的中国区域创新创业指数中, 各城市新增企业注册地址的权重占比达20%; 企查查、天眼查等企业大数据平台以注册地址作为企业位置并划分其所属地区, 而这些按注册地址统计的数据在学术研究得到广泛运用^[2-5]。与

此同时, 许多学者也依据注册地址明确研究区内的企业样本和进行相关数据的处理。如Holguin-Veras等^[6]、Bodenmann等^[7]、Morkutė等^[8]、Rossi等^[9]分别通过注册地址的变更记录来研究美国、瑞士、荷兰和波兰的企业迁移及其影响因素。但是, 企业实体与注册地址的分离现象在一些研究中得到证实。如郭富青^[10-11]将企业实体与注册地址的分离称为注册地址的虚设, 认为该现象已经成为现代企业的趋势, 并提到美国90%以上的公司成立于免税的特拉华州, 却在在该州没有业务办公室; 汪雨卉^[12]在分析上海市初创企业的集聚演化特征时发现53.28%的初创企业实际办公地址与注册地址不一致。此外, 众多离岸企业本质上属于实体位置与注册地址的跨国分离。毫无疑问, 企业实际位置不在注册地会造成按注册地址统计的数据存在误差。

收稿日期: 2023-04-27; 修订日期: 2023-06-24。

基金项目: 国家自然科学基金项目(42171171, 42361050)。[Foundation: National Natural Science Foundation of China, No. 42171171 and 42361050.]

第一作者简介: 胡国建(1992—), 男, 江西九江人, 讲师, 主要研究方向为经济地理与区域发展。

E-mail: guojianhu1992@163.com

*通信作者简介: 钟业喜(1973—), 男, 江西赣州人, 教授, 研究方向为经济地理与空间规划。E-mail: zhongyexi@jxnu.edu.cn

引用格式: 胡国建, 陆玉麒, 钟业喜. 企业区位视角下的数据误差研究: 以A股上市公司为例[J]. 地理科学进展, 2023, 42(12): 2309-2323. [Hu Guojian, Lu Yuqi, Zhong Yexi. Errors in statistical data from the perspective of corporate location: A case study of A-share listed companies. Progress in Geography, 2023, 42(12): 2309-2323.] DOI: 10.18306/dlkxjz.2023.12.003

例如,按注册地址统计的2014年爱尔兰对外直接投资比2007年大幅增加了6000亿美元(这是一个明显偏离于实际的统计值,2014年爱尔兰GDP仅为2591.71亿美元),原因就是期间许多总部位于其他国家的跨国公司将注册地址变更至作为避税天堂的爱尔兰,以减轻税负压力^[13]。

企业实体与注册地址空间分离现象对按后者统计的各类数据所造成的误差影响没有引起学界的重视。现有经济学和统计学关于数据统计误差的研究往往将误差原因归纳为瞒报、数据篡改、抽样误差、统计报表设计不合理等^[14],相关实证研究则是通过各种假设和模型对官方数据(主要是GDP数据)进行质量评估。如选择夜间灯光数据、全社会用电量、税收总额、城镇居民人均可支配收入等与GDP数据具有相关性的指标,通过指数法、相关分析法、生产函数法等方法对GDP数据进行拟合和误差估算^[15-17]。但官方数据的质量评估结果是基于假设条件和定量方法的估计值,并非数据实际的误差情况,这也是为什么国内外研究关于中国GDP规模的评估结果不尽相同的原因^[18]。

从企业实体与注册地址空间分离视角出发,为测度数据实际误差提供可能,即设按企业实际位置统计的数据为正确值,以此对比研究各种数据按注册地址统计的误差。因此,本文在中国A股上市公司总部与注册地址存在跨省分离现象的背景下,分析上市公司在省级尺度上按注册地址统计2019年

各类数据较实际值的偏差程度,具体数据包括各省公司数量、各省公司指标、公司发展环境指标(地区指标)和公司到会计师事务所的地理距离。希望本研究引起经济地理学重视企业实体位置与注册地址空间分离现象及其对数据统计的影响,认识到注册地址在界定企业位置方面的局限性。

1 研究背景与设计

1.1 研究背景

1.1.1 数据统计的误差机制

按注册地址统计企业相关数据是否会导致统计值偏离于现实,需结合企业实体与注册地址分离尺度和统计单元尺度综合判断。若前者尺度超过后者,即企业实体和注册地址中仅一方在统计单元,则会造成企业归属到错误地区并导致统计的数据存在误差。图1a中企业a~g的实体位置与注册地址均分布在不同地区,实际上,A区仅有1家企业(c),B区有4家企业(a、b、e、f),C区有2家企业(d、g),但按注册地址统计会显示A区有3家企业(a、b、d),B区仅有1家企业(g),C区有3家企业(c、e、f)。不仅如此,将企业a~g的各项指标归并到地区或者统计企业a~g的发展环境指标时均会出现“张冠李戴”的情况。若计算企业间的距离,企业a和e实际上在同地区内且距离较近,但按注册地址统计会显示两者在不同地区且距离较远。

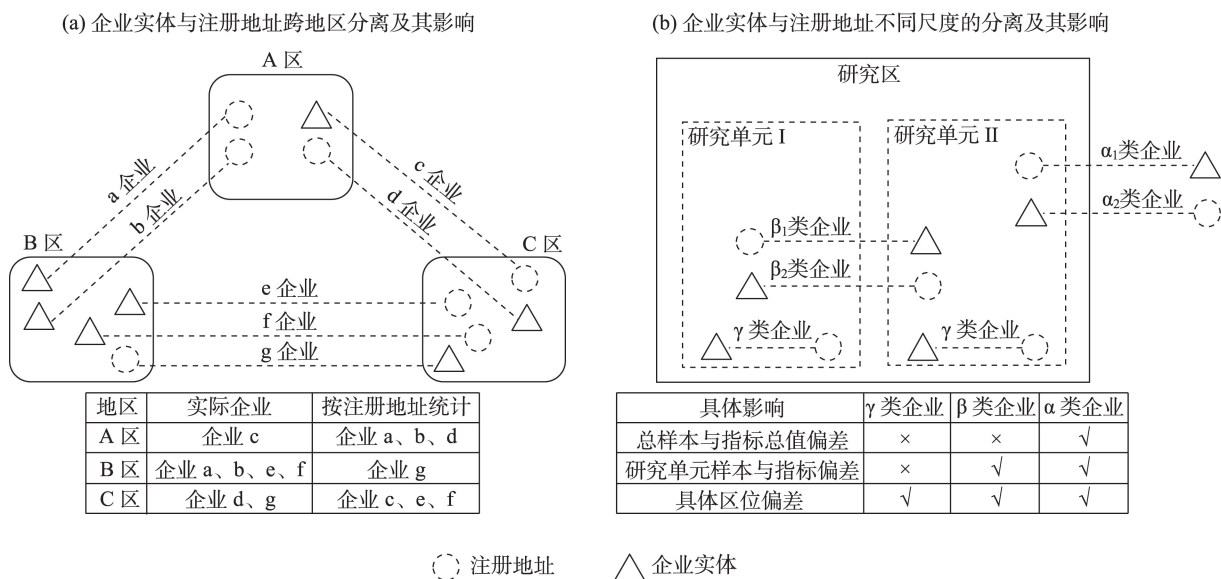


图1 按注册地址统计企业相关数据的误差机制

Fig.1 Mechanism of errors in enterprise statistical data by registered address

在学术研究中,样本企业可能存在实体位置与注册地址多尺度分离。而企业相关的研究可分为两类:一类是研究区内不细分研究单元,如对单个城市的研究;另一类是研究区内细分研究单元,如对长三角地区(研究区)各个城市(研究单元)的研究。前者按注册地址进行数据统计的误差机制相当于后者的简化,因此本文仅以后者为例进行梳理(图1b): γ 、 β 和 α 分别是实体位置与注册地址在研究单元内部分离、跨研究单元分离和跨研究区分离的企业类,三类企业的注册地址均与实体位置存在偏差,偏差程度一般是 α 类企业 $>$ β 类企业 $>$ γ 类企业。基于注册地址判断 β 和 α 类企业位置还会造成研究单元的企业样本以及企业各类数据的统计值偏离于实际(γ 类企业则不会),其中, β_1 类企业实际属于研究单元II但被归到研究单元I, β_2 类企业实际属于研究单元I但被归到研究单元II。此外, α 类企业还会导致研究区的企业总样本及企业相关数据的统计值出现误差(β 类企业则不会),其中 α_1 类企业属于其他地区但被纳入了研究, α_2 类企业属于研究区但在研究中被剔除。

1.1.2 上市公司(总部)地理位置界定的分歧

从企业实体与注册地址空间分离视角量化数据统计误差的关键在于获取到企业的这两种位置信息,公布办公地址与注册地址的上市公司则成为了突破口。对于上市公司等规模较大、多地经营的企业,其地理位置通常被默认为总部所在地^[19-20],而办公地址和注册地址均是国内外文献中上市公司(总部)位置的常见标准。但两者在地理空间上并不完全重合^[21-23],这意味着部分文献中上市公司(总部)位置一定存在偏差,并导致研究数据和结论可能出错。

在此背景下,注册地址和办公地址哪个更能表征总部的地理位置是学界应重视的问题。目前上市公司(总部)地理位置的分歧并没有实证研究的澄清,只有极少数相关文献在数据处理时解释其选择办公地址或注册地址的理由。如刘磊^[20]的研究提到:“因为公司在选择注册地址时往往会考虑税收、政府优惠等因素,但公司的经营业务并不一定在注册地址开展。因此公司的办公地址作为公司与外界联系和业务交流的地方,更具有实际意义”;沈甜甜等^[24]提到:“考虑到现实中存在公司的注册地址

并非经营地的情况,本文在选取地理位置指标时采用上市公司的办公地址所在城市度量”。与之相反,曹洁^[25]提到:“考虑到办公地址的多变性,本文采用两者(基金公司和银行)的注册地址作为两者的所在地”。本文认为,从实际地理空间和经济属性的角度出发,选择办公地址作为上市公司(总部)地理位置更具有合理性,而前期针对办公地址与注册地址跨省分离的上市公司所进行的查证(以年报和官网为主)也显示几乎所有公司的总部都位于办公地址而非注册地址。

所以,本文以办公地址作为上市公司总部的地理位置,并将在此基础上统计的各类数据设为正确值,对比分析按注册地址统计的数据所存在的误差情况。需要说明的是,上市公司仅占全体企业的一小部分,特别是在为数众多的小微企业方面的代表性存在一定局限性。不过,上市公司在国民经济发展中具有重要地位和影响,是各方数据统计的重要对象,也是学界最常用的企业数据来源之一^[26]。因此,本文研究对审视现有众多使用了注册地址作为上市公司(总部)位置的统计数据和相关研究,以及未来相关学者如何处理上市公司相关数据都具有重要的参考价值。且目前来看,只有上市公司可用于量化按注册地址统计数据的误差,因为大部分非上市公司除了注册地址外没有其他的位置信息(除非实地走访^①)。

1.2 研究设计

1.2.1 上市公司样本处理和描述

由于兼顾所有统计尺度不便于分析,本文对公司数量统计、公司指标归并和公司发展环境指标的统计都以省级行政区为单元,而地理距离的误差分析也聚焦于总部与注册地址跨省分离的上市公司。通过Wind金融数据库获取截至2020年1月1日的3743家A股上市公司的总部位置与注册地址且归并到省级行政单元,剔除有两个总部且分布在不同省份的35家公司,剩余3708家公司。结果显示:总部与注册地址跨省分离的公司有233家,正是这些公司造成各省按注册地址统计的上市公司相关数据与实际值存在偏差,表1是233家公司总部与注册地址所在省份的对应关系。可以发现,总部倾向于集中在少数发达地区,62.66%的总部位于京、沪,86.27%的总部位于京、沪、粤、苏、浙,其他省

① 中国经济普查是通过实地走访的方式统计数据,并要求注册地与实际经营地不一致的企业填报其注册地址。如果能获得中国经济普查足够详细的微观数据,对研究企业实体与注册地址的分离现象及其对数据统计误差的影响有很大帮助。

表1 223家公司总部与注册地址的省份分布

Tab.1 Distribution of the office addresses and registered addresses of 233 companies by province

总部 所在 省份	注册地所在省份																															
	浙江	江苏	湖北	山东	辽宁	天津	黑龙江	西藏	湖南	广东	吉林	河北	新疆	广西	安徽	福建	上海	四川	甘肃	山西	重庆	海南	陕西	内蒙古	江西	河南	云南	北京	贵州	青海	宁夏	合计
北京	8	5	5	6	4	9	7	1	2	6	4	6	4	1	1		4	2	1	2	2		2	4	3	1	1		1			92
上海	15	11	1	2	3	1	2	2	2	2			2		2	4					2	2									1	54
广东	4	1	2		3			3	3		2	1		2	3	1		1	1	2	1	1				2						33
江苏	1		2	3	1				1		1						1								1			1	1			13
浙江		1	1								1			2				1	2									1				9
四川								2					1						1				1									5
福建			2				1																1									4
江西		1													1					1												3
河北						1		1						1																		3
陕西				1																1												2
天津																		1			1											2
山东					1													1														2
湖北				1																							1					2
新疆	1																	1														2
重庆		1																											1			2
湖南																						1										1
辽宁					1																											1
安徽	1																															1
河南									1																							1
云南																	1															1
合计	30	20	14	13	12	11	10	9	9	8	8	7	7	6	6	6	6	6	6	6	6	4	4	4	4	3	2	2	2	2	1	233

份的总部均不超过5家。而注册地址则分散在除港澳台外的全部31个省份,其中浙、苏、鄂、鲁、辽、津、黑的注册地址均在10家及以上,合计占比47.21%。

总部与注册地址的空间分布差异与已有文献中的结论相吻合^[22-23,27-28]。即总部等企业实体与注册地址在区位因子方面存在显著区别,前者主要受资源禀赋、劳动力、集聚效应、市场、基础设施等经济类因素的影响,因此倾向于分布在经济发达地区;而注册地址主要受到以税收政策为主的制度类因子的影响,与企业实体相比,更多地分布在具有税收优惠的欠发达地区。例如,宏观尺度上,许多跨国公司注册在经济落后、人口稀少的避税天堂,而总部和经营地多位于世界主要经济体^[28];微观尺度上,实体位置与注册地址市内分离的企业往往将前者设在中心城区,后者则设在相对落后的经济开发区和城市边缘区^[12]。

1.2.2 指标数据与处理

综合来看,注册地址在学术研究中的用途主要包括明确研究区内的企业样本、将企业指标归并到地区,统计企业发展环境指标和计算企业相关的地

理距离。因此数据统计误差的测度从这4个方面进行,具体的数据指标综合考虑了数据的可获得性和指标在学术研究中被采用的广泛程度:

(1) 上市公司数量(样本)偏差。企业空间分布的研究是经济地理学重要的研究命题,其中不少研究以注册地址明确各地区的企业样本或数量。如李汉青等^[29]在研究珠三角制造业集聚特征时通过注册地址识别珠三角各城市的企业样本。因此本文分析各省按注册地址统计的上市公司数量较实际值的误差。

(2) 上市公司指标归并误差。许多研究将企业相关指标归并到注册地,如欧阳艳艳等^[30]在企业投资对城市空气污染影响的研究中将公司对外直接投资金额和专利数归并到注册地址所在城市。因此本文通过Wind金融数据库获得2019年A股上市公司相关数据进行指标归并误差分析。需要强调的是,上市公司往往是多地经营的企业,直接将其各类生产经营数据归并到总部所在地可能与实际情况不符。为此,本文借鉴Coppola等^[31]对国际投资数据统计以及丰田汽车北美公司相关数据处理

中企业位置标准的反思,选取具有流动性且可被总部控制或变现的指标(包括注册资金、市值、总资产、营业收入、营业利润、研发投入),分别以总部和注册地址为依据归并到各个省份,对比分析按注册地址统计的数值偏差程度。相比之下,员工数等不易流动和变现,以及政府补贴等形式复杂和使用受限的指标不纳入考虑。

(3)上市公司发展环境指标的统计误差。除了按注册地址归并企业指标,同样有不少研究以注册地址为依据统计企业的发展环境指标,作为定量模型的解释变量。如王志高等^[32]在企业创新成功率与城市规模间关系的研究中统计了企业注册地址所在城市的人口数、GDP、客运总量、货运总量、国际互联网用户数、实际使用外资等诸多指标。因此本文通过 Easy Professional Superior 数据平台获取2019年各省人均GDP、居民人均可支配收入、建成区面积、城市绿地面积、规模以上工业企业数量、外商投资企业登记数量、医院数、公园数8个指标,分析按注册地址统计上市公司发展环境指标较实际值存在的误差。

(4)上市公司相关距离的计算误差。距离是地理学的核心概念^[33],以注册地址作为企业位置计算相关距离是学术研究中常见的做法。例如一些研究在探讨会计师事务所对上市公司审计事项的影响时根据注册地址计算公司到会计师事务所的地理距离^[34-37]。因此,本文以上市公司聘用的会计师事务所为例,分析按注册地址计算公司到会计师事务所的距离较实际值的偏差程度。需要强调的是,总部与注册地址在任何尺度的分离都会造成按注册地址统计的距离指标出现误差,因此,本文中距离误差的分析仅聚焦于233家总部与注册地址跨省分离的公司,不考虑这些公司对全体公司到会计师事务所距离统计的整体误差影响。上市公司聘用的会计师事务所及其地址由 Wind 金融数据库提供,使用谷歌地图地址解析服务得到会计师事务所以及上市公司总部、注册地址的经纬度,用于距离的计算。

1.3 按注册地址统计的误差率

将按上市公司总部位置进行统计的指标值设为正确值,则按注册地址进行统计的指标误差率为:

$$E_i = \frac{|x_{i\text{注}} - x_{i\text{总}}|}{x_{i\text{总}}} \times 100\% \quad (1)$$

式中: E_i 为*i*指标按注册地址统计的误差率,也称相对误差; $x_{i\text{注}}$ 和 $x_{i\text{总}}$ 分别是*i*指标按注册地址和总部统计的数值,两者的差值为绝对误差。

2 以注册地址为公司位置的数据统计误差结果分析

2.1 公司样本偏差分析

总部与注册地址同省的3475家公司在省份归属上无争议,但两者跨省分离的233家公司会造成按注册地址统计的各省公司数量存在误差,且一家公司会影响两个省份,即总部所在省份漏算而注册地址所在省份多算。图2是各省拥有总部与注册地址均在省内、仅总部在省内和仅注册地址在省内的3类公司数量。

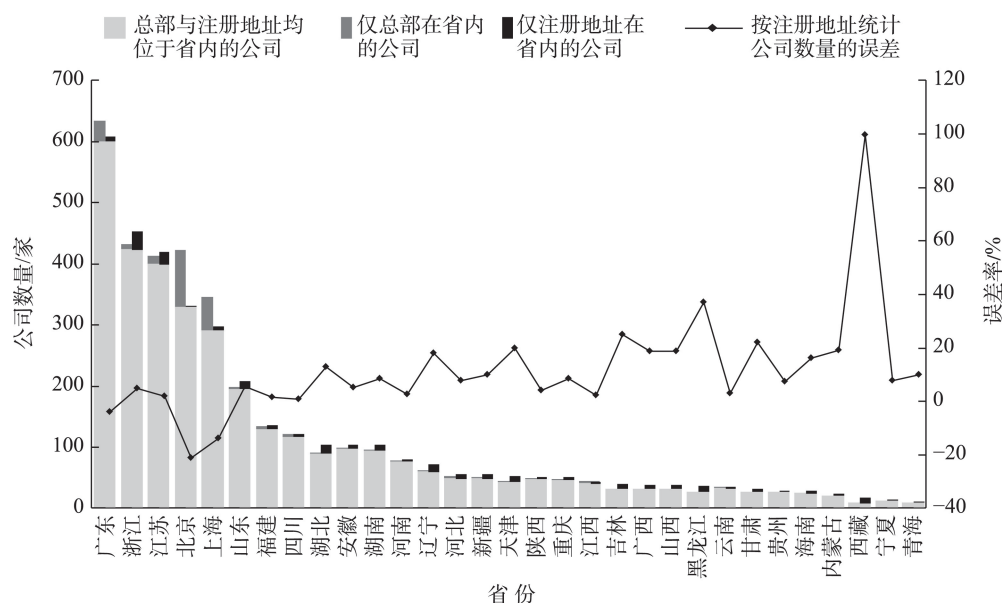
按注册地址统计的各省上市公司数量误差取决于总部数和注册地址数的差值。其中,北京、上海和广东按注册地址统计的公司数量分别由实际数量的422家、346家、634家降至332家、298家、609家,降幅分别为21.33%、13.87%、3.94%。而其他省份的注册地址数均多于总部数,即按注册地址统计会高估公司数量,按误差率可分为4个级别:

①四川、福建、江苏、江西、河南、云南、陕西和浙江按注册地址统计公司数量的误差率较小,较实际数量分别提升0.82%、1.49%、1.69%、2.27%、2.56%、2.86%、4.00%、4.85%。其中江苏和浙江的注册地址比总部分别多7家和21家,但两省拥有众多总部和注册地址均在省内的公司,导致误差率小,而其他6省的注册地址数与总部数相差均不超过2家。

②安徽、山东、贵州、河北、宁夏、重庆、湖南、新疆和青海按注册地址统计公司数量的误差率一般,较实际值分别提升5.05%、5.56%、7.41%、7.69%、7.69%、8.33%、8.33%、9.80%和10.00%。不过,这9个省份的注册地址数和总部数的平均差值(4.56家)与第1类省份(4.63家)相比反而微降,误差率更高的原因是各省拥有总部与注册地址均在省内的公司平均数量(64.78家)较第1类省份(158.88家)大幅下降。

③湖北、海南、辽宁、广西、山西、内蒙古按注册地址统计公司数量的误差率较高,较实际值分别提升13.04%、16.00%、18.03%、18.75%、18.75%和19.05%,它们的注册地址数比总部数平均多7.17家。

④天津、甘肃、吉林、黑龙江和西藏按注册地址统计公司数量的误差率高,较实际值分别提升



注:各省按总部(注册地址)统计的公司数量=总部与注册地址均在省内的公司数量+仅总部(注册地址)在省内的公司数量。

图2 按总部与注册地统计的各省A股上市公司数量

Fig.2 The number of listed A-share companies in each province by headquarter and place of registration

20.00%、22.22%、25.00%、37.04%和100.00%。这5省的注册地址数比总部数平均多8.40家,其中,总部与注册地址都在西藏的公司和仅注册地址在西藏的公司各有9家。

可以发现,尽管之前使用注册地址分析中国上市公司空间格局的研究已经得出上市公司分布不均衡的结论^[38-39],但中国上市公司实际的不均衡程度比注册地址显示的更加严重。原本就拥有众多上市公司的北京、上海和广东三地吸引了大量与注册地址跨省分离的总部,其中北京和上海还对周边地区的上市公司总部产生明显的虹吸效应,如有26家注册在浙江和江苏的上市公司选择将总部设在上海。而经济相对不发达的中西部和东北地区的省份基本不存在与注册地址跨省分离的总部,却拥有不少与总部跨省分离的注册地址,致使其真实的上市公司数量少于注册地址数量。

2.2 公司指标归并误差分析

公司指标包括注册资金、市值、总资产、营业收入、营业利润、研发投入,6项指标按注册地址和总部统计的结果如图3所示。北京由于存在许多规模庞大的央企,按两种位置标准统计的各项指标值均位列第一。所有省份根据各指标按注册地址统计的整体误差率可分为以下5类:

(1) 按注册地址统计的指标值整体低于实际值:包括北京、上海、广东和江西,各项指标较实际

值的平均降幅分别为6.54%、11.02%、0.58%和3.96%。其中,前三者的指标值下降是因为按注册地址统计的公司数量减少。不过北京各指标的平均降幅显著小于公司数量的降幅,主要原因是许多大型央企成为北京指标统计的“压舱石”,仅总部在北京的92家公司造成的指标误差程度有限。江西按注册地址统计的各指标整体低于实际值是因为仅注册地址在该省的4家公司整体经营不佳,各项指标值显著低于仅总部在江西的3家公司。

(2) 按注册地址统计的指标值整体略高于实际值(各指标增幅多在5.00%及以内):包括四川、陕西、福建、河南、宁夏、云南、山西、贵州、浙江。其中,除山西外的8个省份指标误差率较小是因为注册地址数与总部数较为接近,而山西的注册地址数比总部数多18.75%,但由于仅注册地址在山西的6家公司规模较小,使得按注册地址统计的各项指标平均仅上升2.62%。另外,仅注册地址在宁夏的某公司2019年亏损23.02亿元,造成宁夏的营业利润按注册地址统计大幅下降65.73%。

(3) 按注册地址统计的指标值整体增幅一般(各指标增幅多为5.01%~15.00%):包括湖南、内蒙古、江苏、安徽、山东、重庆、吉林、海南。其中,江苏的注册地址仅比总部多1.69%,但仅注册地址在江苏的公司各项指标数值显著高于仅总部在江苏的公司,因此按注册地址统计的各项指标值较实际值平

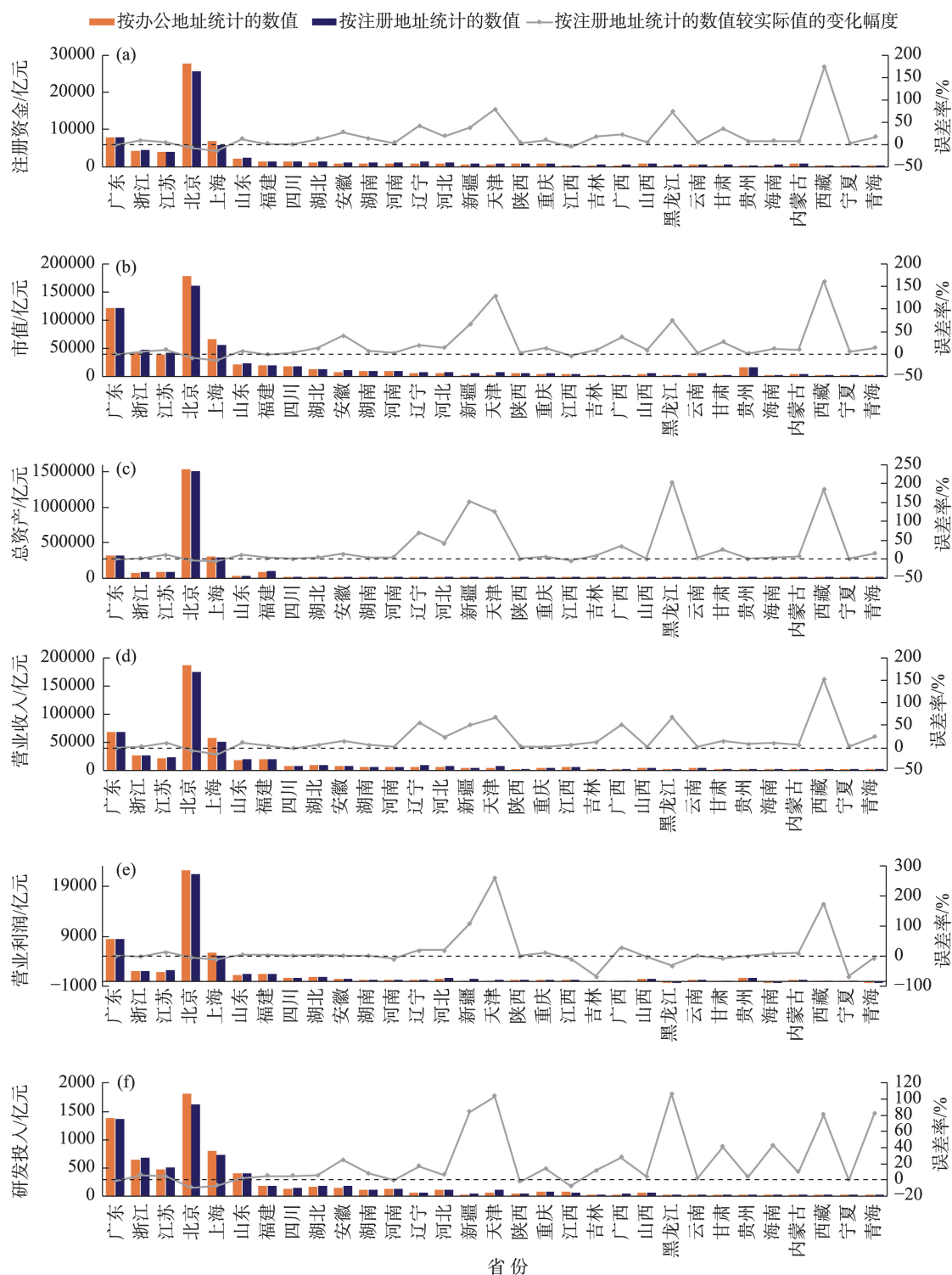


图3 按总部与注册地址统计的各省上市公司指标及按注册地址统计的误差率

Fig.3 Indicators of listed companies in various provinces by headquarters and registered address and error rate by registered address

均上升9.14%。仅注册地址在吉林的8家公司共亏损52.18亿元,导致吉林按注册地址统计的营业利润较实际值下降66.35%。另外,仅注册地址在海南的某家高技术制造公司投入了大量研发费用,使得

按注册地址统计海南的研发费用指标增长42.55%。

(4) 按注册地址统计的指标值整体增幅较大(各指标增幅多为15.01%~40.00%):包括河北、湖北、甘肃、青海、广西、辽宁。不过,仅注册地址在湖北、青

海和甘肃3个省份的上市公司存在不少亏损的情况,导致湖北按注册地址统计的营业利润较实际值仅提升3.56%,青海和甘肃更是分别下降6.15%和5.50%。

(5) 按注册地址统计的指标值整体增幅极大(各指标增幅多大于40.00%):包括黑龙江、新疆、天津和西藏,指标平均增幅分别为82.14%、83.37%、127.71%、154.81%。其中,新疆和天津的注册地址数仅分别比总部数多9.80%和20.00%,但由于仅注册地址在新疆的2家国有公司,以及仅注册地址在天津的8家央企的各项指标值显著高于上市公司的平均水平,导致两地按注册地址统计的各项指标值较实际值大幅上升。

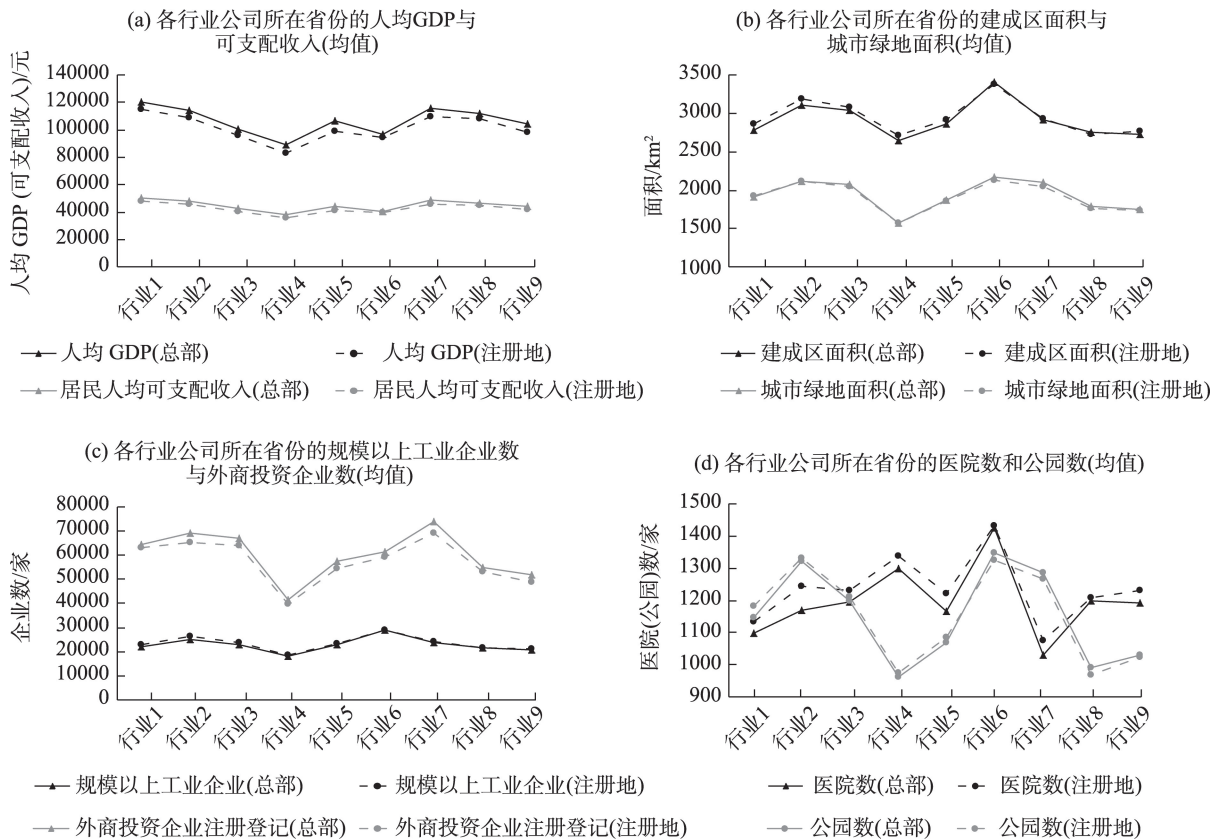
整体而言,各省份按注册地址统计上市公司相关指标的误差情况很大程度上取决于公司数量的误差状况。不过由于上市公司规模、经营状况之间的显著差异,使得一些省份按注册地址统计的指标误差情况比公司数量误差更严重。特别是甘肃、青

海、广西、辽宁、新疆、黑龙江、西藏等东北、西北和西南地区的省份,总部与注册地址都在这些地区的上市公司数量较少,且一般规模小、经营状况欠佳,因此经常出现个别公司的错误计算会造成指标大幅变化的情况(尤其是营业利润)。总体上,全部186个统计值(31个省份×6项指标)的平均误差率为24.84%,可以认为按注册地址归并上市公司的指标存在较大误差。

2.3 公司发展环境指标的统计误差分析

本小节中,企业发展环境指标是被统计和归并的对象,因此误差分析按上市公司的行业进行归类。具体将3708家上市公司分为9大行业(图4)。依据总部和注册地址分别统计各行业公司所在省份的人均GDP、居民人均可支配收入、建成区面积、城市绿地面积、规模以上工业企业数、外商投资企业注册登记数、医院数、公园数8个指标的均值(图4)。可以发现:

(1) 按注册地址统计上市公司所在省份的人均



注:行业1为信息传输、软件和信息技术服务业,行业2为建筑业,行业3为交通运输、仓储和邮政业,行业4为电力、热力、燃气及水生产和供应业,行业5为批发和零售业,行业6为制造业,行业7为房地产业,行业8为金融业,行业9为其他行业。

图4 按总部与注册地址统计的各上市公司所在省份的指标

Fig.4 Indicators of provinces where listed companies are located by headquarters and registered address

GDP、人均可支配收入和外商投资企业注册数较实际值出现明显下降,各行业平均降幅分别为5.08%、4.72%和4.47%。主要原因是总部与注册地址跨省分离的公司中86.27%的总部位于北京、上海、广东、浙江和江苏(表1),这5个省份的人均GDP、人均可支配收入和外商投资企业注册登记数在全国排名靠前,造成按注册地址统计的指标值明显低于实际值。

(2) 按注册地址统计上市公司所在省份的城市绿地面积、公园个数和建成区面积与实际值较为接近,各行业的平均变化幅度分别为-0.86%、0.11%和1.23%。主要原因是北京和上海的城市绿地面积、公园数和建成区面积3个指标在全国排名中等(略低于各省平均值),即总部在北京和上海而注册地址在其他省份的众多公司没有造成按注册地址统计的数值出现明显偏差。

(3) 按注册地址统计上市公司所在省份的规模以上工业企业数和医院数较实际值出现明显提升,各行业分别增加2.34%和3.23%。主要原因是北京和上海拥有规模以上工业企业数量、医院数量在所有省份中排名落后,因此总部在北京和上海而注册地址位于其他省份的众多公司造成按注册地址统计的数值整体高于实际值。

因此,按注册地址统计上市公司的发展环境指标会造成一定偏差,尽管72个统计值(9个行业×8项指标)平均误差率仅为3.04%,但指标之间、行业之间的差异明显。之所以公司发展环境指标的平均误差率明显小于各省公司指标归并的平均误差率,一是因为前者统计的是各地区的均值,后者统计的是公司指标的和值,二是因为各地环境指标间的差异显著小于公司指标间的差异。整体而言,由于总部与注册地址跨省分离的公司绝大部分将总部设在北京、上海、广东、浙江和江苏,各项指标按注册地址统计的误差率很大程度上取决于这5个省份的情况,即这5个省份的指标值排名靠前会导致按注册地址统计的数值低于实际值,反之会高于实际值。唯一例外的是制造业(行业6)公司所在省份的规模以上工业企业数按注册地址统计数值较实际值下降0.73%,其他行业则高于实际值,因为与注册地址跨省分离的制造业公司总部中有46.90%位于北京和上海(即受两市的影响相对较小),还有28.32%位于规模以上工业企业数最多的两个省——广东与江苏。

2.4 地理距离的计算误差分析

总部和注册地址跨省分离的233家公司中分别有134、49、20和17家公司的会计师事务所位于北京、上海、浙江和湖北,其他13家公司的会计师事务所分散在天津、江苏、广东、福建、山东和江西。图5是各公司总部与注册地址分别到会计师事务所的距离,可以发现图中有3个集中区:①有72家公司总部到审计机构距离小于50 km,以总部和会计师事务所都在北京或上海的公司为主(69家);②有70家公司的总部到审计机构距离介于1040~1140 km之间,以总部在北京而会计师事务所在上海的公司(25家)和总部在上海而会计师事务所在北京的公司(23家)为主;③有22家公司的总部到审计机构距离介于1890~1990 km之间,以总部在广东而会计师事务所在北京的公司为主(21家)。在这3个值域的公司合计占比达70.39%。

图5中的公司越接近 $y=x$ 线而远离横纵坐标轴,按注册地址统计的距离误差程度越小,反之越大。233家公司按注册地址计算到会计师事务所距离的平均误差为739.94 km。其中有146家公司到会计师事务所的距离会被高估,高估幅度在0~200、200~1000、1000~2000及2000 km以上的公司分别有36、54、47和9家,平均高估827.18 km;另有87家公司到会计师事务所的距离会被低估,低估幅度在0~200、200~1000、1000~2000及2000 km以上的公司分别有26、46、14和1家,87家公司被平均低估593.52 km。之所以距离被高估的公司多于被低估的公司,以及被高估的平均距离大于被低估的平均距离,主要是因为75家公司的总部和会计师事务所同省(主要在北京和上海),平均相距仅17.85 km;相比之下,注册地址和会计师事务所同省的公司仅17家,且平均相距167.90 km。

综上,不同的省份、行业和指标按注册地址统计数据的误差率差异显著,不少统计项的误差率超过了5%(即统计学者一般认为经济数据可信的误差率阈值^[40]),整体误差幅度不容忽视。而且企业实体与注册地址分离视角下的数据统计误差问题涉及多种分离尺度、多种统计单元尺度和多种数据类型,本文的研究结果并没有完全呈现数据按注册地址统计的误差情况。一方面,公司数量、公司指标、公司发展环境指标、地理距离等可量化的数据指标按注册地址统计的误差程度要高于上述分析中呈现的结果。例如,仅注册地址在江苏、仅总部在江

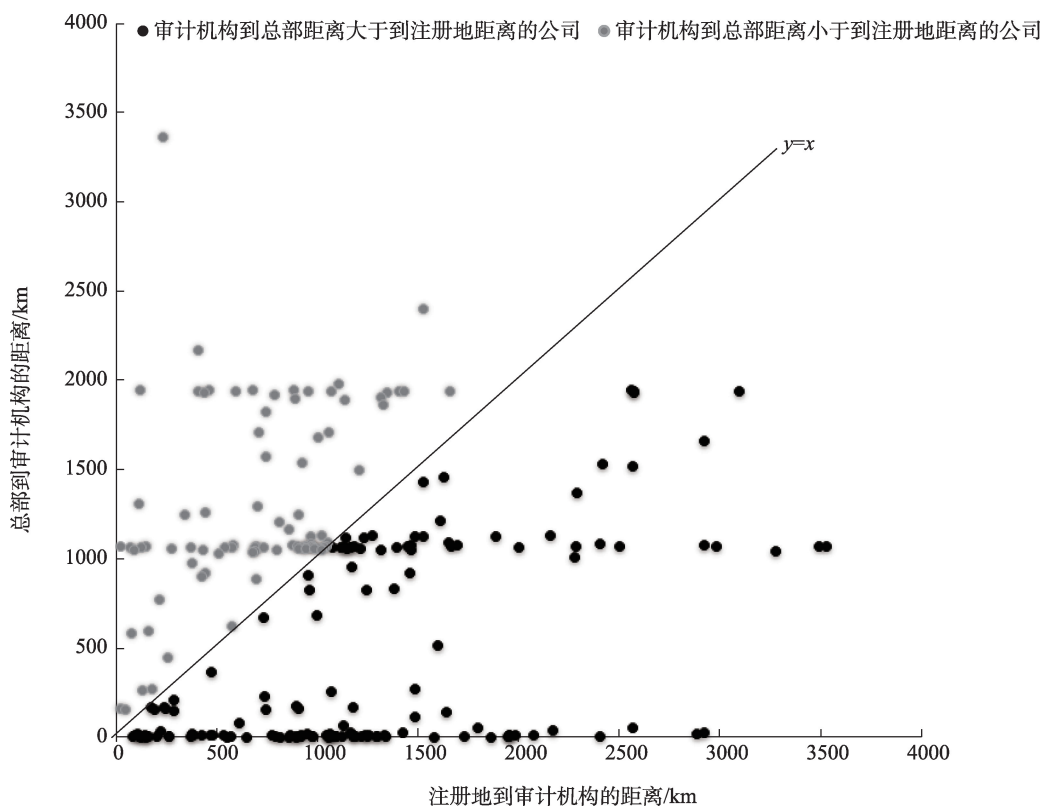


图5 233家上市公司总部和注册地址到会计师事务所的距离

Fig.5 Distances from headquarters and registered addresses to accounting firms of 233 listed companies

苏和两者均在江苏的公司分别有20、13和400家,因此按注册地址统计的公司数量误差值和误差率分别为7家和1.69%,但实际上有33家公司被错误归并。即某省按注册地址统计的公司数量误差是仅注册地址在省内和仅总部在省内的两类公司数量之差,但实际被错误归并的公司数量是两类公司之和。当研究中涉及非量化的企业数据时(如企业迁移的起点和终点),按注册地址进行统计的误差会全部呈现。另一方面,企业实体与注册地址大尺度上的分离会影响小尺度上的数据统计,而小尺度上的分离一般不影响大尺度上的数据统计(图1b)。本文数据误差的统计尺度为省级行政区,城市和区/县尺度上的各类数据按注册地址统计的误差情况预计会更严重。

3 结论和讨论

3.1 结论

许多以注册地址作为企业位置进行统计的社会经济数据会因企业实体与注册地址空间分离现象而偏离于现实。目前学界对于这一重要问题缺

乏重视,因此,本文分析了2019年A股上市公司按注册地址统计各类数据较实际值存在的偏差。主要结论如下:

(1) 总部与注册地址跨省分离的233家A股上市公司大多将总部设在北京、上海、广东等经济发达地区,而注册地址分散在除港澳台外的所有省份。因此,按注册地址统计北京、上海和广东的公司数量低于实际值,其他省份则高于实际值,这意味着上市公司实际的地域分布比注册地址显示的更加不均衡。各省按注册地址统计的公司数量误差率平均为14.23%,误差率的大小更多取决于总部与注册地址都在省内的公司数量(与误差率成反比例关系),而非注册地址数与总部数的差值。

(2) 各省按注册地址统计上市公司6项指标的误差率平均为24.84%,误差幅度较大。北京、上海、广东和江西统计的公司指标整体较实际值下降,其他省份按注册地址统计的指标值大多高于实际值,其中各指标的平均高估幅度在0~5.00%、5.01%~15.00%、15.01%~40.00%和40.00%以上的分别有9、8、6和4个省份。总体而言,各省的上市公司指标归并误差很大程度上由公司数量的误差决定,不过

指标归并误差的幅度大于公司数量误差幅度,特别是上市公司数量较少的东北、西北和西南等地的一些省份,存在个别公司的错误计算导致指标发生大幅波动的情况。

(3) 各行业上市公司按注册地址统计公司发展环境的8个指标平均误差率为3.04%,由于总部与注册地址跨省分离的公司中总部大多位于京、沪、粤、浙、苏,各指标的误差率很大程度上取决于这5个省份(尤其是北京和上海)的情况。即这5个省份的指标值较高会导致按注册地址统计的数值低于实际值,如人均GDP、人均可支配收入和外商投资企业注册数;反之会高于实际值,如规模以上工业企业数和医院数。

(4) 按注册地址计算公司到会计师事务所距离的误差取决于总部、注册地址和会计师事务所的位置关系。由于不少公司在总部所在地聘用会计师事务所,总部和注册地址跨省分离的233家公司中按注册地址统计有146家公司到会计师事务所的距离被高估,平均高估827.18 km;相比之下,仅有87家公司到会计师事务所的距离被低估,平均低估593.52 km。

3.2 讨论

企业实体与注册地址空间分离导致的数据统计误差问题预计将长期存在。一方面,大部分国家的商事登记制度仅要求企业登记注册地址,即数据统计中没有替代注册地址的其他位置标准,且未来商事登记制度改革中加入经营场所作为登记事项无论是在法理上还是改革方向上都不具有可行性^[10-11,27]。另一方面,企业实体与注册地址的空间分离现象很难得到有效控制。即使在原则上禁止两者不一致的中国,也因为基层监管力度和惩戒力度的不足、相关企业容易通过虚设办公地址规避监管、部分地区为争夺税源和完成招商引资考核放任甚至鼓励在外地经营的企业仅将注册地址迁入辖区内等一系列原因,造成企业实体与注册地址的分离现象相当常见。因此,各类社会经济数据按注册地址统计的误差难以根除是由客观统计能力不足和主观造假共同造成,而其中主观造假的责任方主要包括企业和统计机构。事实上,国际清算银行(BIS)、国际货币基金组织(IMF)、美国联邦储备委员会(FED)都曾倡议在企业真实国籍基础上重新统计各种投资流动数据^[31];美国也在法律政策层面针对跨国公司主要经营地与注册地址的分离现象进行

了长期斗争(主要目的是限制跨国公司在避税天堂注册而造成税基侵蚀),如2004年出台的《美国就业创造法案》(American Jobs Creation Act)和2017年出台《减税和就业法案》(Tax Cut and Jobs Act)都对美国跨国公司在避税天堂注册的行为和动机产生了深远影响^[28,41]。但这些努力对于消除企业实体与注册地址空间分离现象及其负面影响还远远不够。

除了数据误差本身外,数据误差对政府决策、企业战略制定、学术研究等方面的负面影响也需要引起关注。事实上,已经有少数学者在上市公司相关的实证研究中意识到公司(总部)位置与注册地址并不总是重合,并采取了一些措施以保障研究结论的准确性或严谨性。包括:①使用办公地址进行稳健性检验。如谭劲松等^[21]在上市公司政治联系对企业长期借款影响的研究中,以注册地址统计各城市上市公司数量进行实证分析,再依据办公地址统计各城市上市公司数量进行稳健性检验。②通过其他途径获取办公地址与注册地址不一致的上市公司(总部)位置。如文嫣等^[42]提到:“北京电影企业是指注册地址在北京的电影企业,而少数办公地址与注册地址不一致的电影企业则以其官网提供的总部地址为准”。③剔除注册地址与办公地址不一致的公司。如汲昌霖^[43]涉及上市公司制度环境的研究,鉴于办公地址与注册地址不一致的公司可能受到两个地区制度环境的共同作用,剔除了这类公司。相比之下,非上市企业的位置信息大多只有注册地址,因此非上市企业的相关研究即使意识到企业实体与注册地址存在不一致的情况,也只能在数据处理或者研究不足中进行局限性说明,如Dong等^[1]采用注册地址构建了2005—2015年间中国约2550万家企业高精度的网格化数据集,并在研究不足中提到:“采用的企业注册地址可能与运营地址不一样,特别是一些企业可能在实施税收补贴等政策的工业园区内注册,但在其他地方经营”。整体来看,上述意识到注册地址在界定企业位置方面存在局限性的研究还较少,且采取的措施并不能完全消除数据误差的影响,特别是对于为数众多的非上市企业,并没有有效的改进手段。

经济地理学作为一门研究经济活动区位的学科,应承担研究企业实体与注册地址空间分离现象所导致的数据误差及其负面影响的主要任务。一方面,实体位置与注册地址的空间分离本身就是一种企业的区位行为,按注册地址统计的数据误差属

于区位影响的范畴。另一方面,经济地理学除了会受到本文关注的公司数量、公司指标、发展环境指标和地理距离的统计误差影响外,还有关于企业的选址、空间分布、迁移等直接涉及企业具体位置的大量研究可能因为企业实体与注册地址的分离现象而导致数据和结论出错。例如,根据本文研究结果,北京市不仅按注册地址判断的上市公司数量误差率高达21.32%,还有123家上市公司总部与注册地址在北京市内部分离,占实际北京上市公司(按总部算)数量的29.15%。对于许多依据注册地址研究上市公司(总部)空间分布的文献而言,仅北京市的公司样本中存在问题(样本误差和位置偏差)的比例就超过50%。

本文以A股上市公司为例,考察企业相关数据按注册地址统计的误差情况,但研究还存在一些不足:① 本文(以及大部分上市公司相关研究)将总部作为上市公司地理位置的标准,而现实中许多企业的经营活动并非集中于一处。本文在指标归并误差中考虑到了这一局限性,但并不能完全规避多地经营企业指标汇总到总部所在地的统计结果与实际情况之间的差距。事实上,如何界定和捕获多地经营企业(尤其是跨国公司)的地理位置或所属地区一直是学界、政府和国际组织备受争论的议题,因为关系到数据统计、税权分配、企业“国籍”等许多重要事项。② 本文就企业实体与注册地址分离导致的数据误差负面影响的讨论和所举案例集中在学术研究层面,在政府、企业的判断和决策层面的负面影响几乎没有涉及。一是因为相关资料的缺乏;二是因为用于政府、企业决策的各类官方数据的整体误差情况,以及其中有多少是由企业实体与注册地址分离现象导致的尚不明确。

参考文献(References)

- [1] Dong L, Yuan X H, Li M, et al. A gridded establishment dataset as a proxy for economic activity in China [J]. *Scientific Data*, 2021, 8: 5. doi: 10.1038/s41597-020-00792-9.
- [2] 毛刘乐, 盛科荣, 张杰, 等. 金融网络嵌入对中国城市创业活力的影响研究 [J]. *地理科学进展*, 2022, 41(12): 2244-2257. [Liu Le, Sheng Kerong, Zhang Jie, et al. The impact of financial network embedding on entrepreneurial vitality of cities in China. *Progress in Geography*, 2022, 41(12): 2244-2257.]
- [3] 章屹祯, 汪涛, 张晗. 基于金融细分行业的长三角城市网络的组织模式及驱动因素 [J]. *地理科学进展*, 2022, 41(4): 567-581. [Zhang Yizhen, Wang Tao, Zhang Han. Organizational models and driving factors of the Yangtze River Delta urban network based on different financial industries. *Progress in Geography*, 2022, 41(4): 567-581.]
- [4] 赵龙凯, 彭传国. 封闭式基金折价与管理绩效的实证研究 [J]. *金融研究*, 2008(4): 102-121. [Zhao Longkai, Peng Chuanguo. An empirical study of closed-end fund discounts and managerial performance. *Journal of Financial Research*, 2008(4): 102-121.]
- [5] 钱肖颖, 孙斌栋. 基于城际创业投资联系的中国城市网络结构和组织模式 [J]. *地理研究*, 2021, 40(2): 419-430. [Qian Xiaoying, Sun Bindong. The spatial structures and organization patterns of China's city network based on inter-city startup investments. *Geographical Research*, 2021, 40(2): 419-430.]
- [6] Holguin-Veras J, Xu N, Levinson H S, et al. An investigation on the aggregate behavior of firm relocations to New Jersey (1990-1999) and the underlying market elasticities [J]. *Networks and Spatial Economics*, 2005, 5(3): 293-331.
- [7] Bodenmann B R, Axhausen K W. Destination choice for relocating firms: A discrete choice model for the St. Gallen region, Switzerland [J]. *Papers in Regional Science*, 2012, 91(2): 319-341.
- [8] Morkutė G, Koster S. Human capital as a location factor: An empirical investigation of relocating firms and their labour force in the Netherlands [J]. *Papers in Regional Science*, 2018, 97(3): 595-616.
- [9] Rossi F, Dej M. Where do firms relocate? Location optimisation within and between Polish metropolitan areas [J]. *The Annals of Regional Science*, 2020, 64(3): 615-640.
- [10] 郭富青. 我国企业住所与经营场所分离与分制改革的法律探析 [J]. *现代法学*, 2020, 42(2): 145-156. [Guo Fuqing. Legal research on the separation system reform of the enterprises' domicile and the business place in China. *Modern Law Science*, 2020, 42(2): 145-156.]
- [11] 郭富青. 我国企业住所制度及改革的法理检讨与前瞻性构想 [J]. *上海政法学院学报(法治论丛)*, 2020, 35(2): 87-104. [Guo Fuqing. Legal review and perspective idea on enterprise residence system and its reform in China. *Journal of Shanghai University of Political Science & Law (the Rule of Law Forum)*, 2020, 35(2): 87-104.]
- [12] 汪雨卉. 多源大数据视角下上海市初创企业集聚演化特征研究 [D]. 上海: 上海师范大学, 2019. [Wang Yuhui. Study on agglomeration and evolution characteristics of start-ups in Shanghai from the perspective of multi-source big data. Shanghai, China: Shanghai Normal University, 2019.]
- [13] Lane P R, Milesi-Ferretti G M. The external wealth of na-

- tions revisited: International financial integration in the aftermath of the global financial crisis [J]. *IMF Economic Review*, 2018, 66(1): 189-222.
- [14] 陈林. 中国工业统计的理论体系与制度变迁: 兼议中国工业企业数据的部分系统性误差 [J]. *经济科学*, 2019 (4): 69-80. [Chen Lin. The theoretical system and institutional change of China's industrial statistics: And a discussion on systematic errors in China's industrial enterprise database. *Economic Science*, 2019(4): 69-80.]
- [15] Doll C N H, Muller J P, Morley J G. Mapping regional economic activity from night-time light satellite imagery [J]. *Ecological Economics*, 2006, 57(1): 75-92.
- [16] Yue W Z, Gao J B, Yang X C. Estimation of gross domestic product using multi- sensor remote sensing data: A case study in Zhejiang Province, East China [J]. *Remote Sensing*, 2014, 6(8): 7260-7275.
- [17] Fernald J G, Hsu E, Spiegel M M. Is China fudging its GDP figures? Evidence from trading partner data [J]. *Journal of International Money and Finance*, 2020(12): 110-116.
- [18] Plekhanov D. Quality of China's official statistics: A brief review of academic perspectives [J]. *The Copenhagen Journal of Asian Studies*, 2017, 35(1): 76-101.
- [19] 陈冬华, 胡晓莉, 梁上坤, 等. 宗教传统与公司治理 [J]. *经济研究*, 2013(9): 71-84. [Chen Donghua, Hu Xiaoli, Liang Shangkun, et al. Religious tradition and corporate governance. *Economic Research Journal*, 2013(9): 71-84.]
- [20] 刘磊. 我国A股市场地域联动性的实证研究 [D]. 大连: 东北财经大学, 2016. [Liu Lei. An empirical research of co-movement of location evidence from China A-share stock market. Dalian, China: Dongbei University of Finance and Economics, 2016.]
- [21] 谭劲松, 陈艳艳, 谭燕. 地方上市公司数量、经济影响力与企业长期借款: 来自我国A股市场的经验数据 [J]. *中国会计评论*, 2010, 8(1): 31-52. [Tan Jinsong, Chen Yanyan, Tan Yan. The number of regional listed companies economical effect and corporate long-term loan. *China Accounting Review*, 2010, 8(1): 31-52.]
- [22] 潘峰华, 夏亚博, 刘作丽. 区域视角下中国上市企业总部的迁址研究 [J]. *地理学报*, 2013, 68(4): 449-463. [Pan Fenghua, Xia Yabo, Liu Zuoli. The relocation of headquarters of public listed firms in China: A regional perspective study. *Acta Geographica Sinica*, 2013, 68(4): 449-463.]
- [23] 胡国建, 金星星, 陆玉麒, 等. 中国上市公司总部与注册地跨城市分离的格局、形成过程和影响因素 [J]. *地理研究*, 2021, 40(2): 402-418. [Hu Guojian, Jin Xingxing, Lu Yuqi, et al. Study on the separation of listed companies' headquarters from registered cities and their influencing factors. *Geographical Research*, 2021, 40(2): 402-418.]
- [24] 沈甜甜, 汪洋. 上市公司地理位置、风险投资参与与股票流动性: 基于创业板市场的实证研究 [J]. *金融理论与实践*, 2020(8): 85-95. [Shen Tiantian, Wang Yang. Boss company location, venture capital participation and stock liquidity: Empirical research based on the ChiNext market. *Financial Theory & Practice*, 2020(8): 85-95.]
- [25] 曹洁. 金融集团内的信息流动与基金投资业绩 [D]. 马鞍山: 安徽工业大学, 2012. [Cao Jie. Information flows within financial conglomerates and the investment performance of mutual funds. Ma'anshan, China: Anhui University of Technology, 2012.]
- [26] 龙小宁, 林志帆. 中国制造业企业的研发创新: 基本事实、常见误区与合适计量方法讨论 [J]. *中国经济问题*, 2018(2): 114-135. [Long Xiaoning, Lin Zhifan. R&D and innovations of Chinese manufacturing firms: Basic facts, common misunderstandings, and discussions on the appropriate choice of econometric methods. *China Economic Studies*, 2018(2): 114-135.]
- [27] 胡国建, 陆玉麒, 胡舒云. 顾及企业注册地址的区位理论研究 [J]. *地理研究*, 2022, 41(2): 580-595. [Hu Guojian, Lu Yuqi, Hu Shuyun. Research on location theory taking into account the registered address of enterprises. *Geographical Research*, 2022, 41(2): 580-595.]
- [28] Janský P. The costs of tax havens: Evidence from industry-level data [J]. *Applied Economics*, 2020, 52(29): 3204-3218.
- [29] 李汉青, 袁文, 马明清, 等. 珠三角制造业集聚特征及基于增量的演变分析 [J]. *地理科学进展*, 2018, 37(9): 1291-1302. [Li Hanqing, Yuan Wen, Ma Mingqing, et al. Manufacturing industry agglomeration characteristics in the Pearl River Delta and evolution based on growth data. *Progress in Geography*, 2018, 37(9): 1291-1302.]
- [30] 欧阳艳艳, 黄新飞, 钟林明. 企业对外直接投资对母国环境污染的影响: 本地效应与空间溢出 [J]. *中国工业经济*, 2020(2): 98-121. [Ouyang Yanyan, Huang Xinfei, Zhong Linming. The impact of outward foreign direct investment on environmental pollution in home country: Local and spatial spillover effects. *China Industrial Economics*, 2020(2): 98-121.]
- [31] Coppola A, Maggiori M, Neiman B, et al. Redrawing the map of global capital flows: The role of cross-border financing and tax havens [J]. *The Quarterly Journal of Economics*, 2021, 136(3): 1499-1556.
- [32] 王志高, 王如玉, 梁琦. 企业创新成功率与城市规模 [J].

- 统计研究, 2016, 33(7): 55-63. [Wang Zhigao, Wang Ruyu, Liang Qi. The success rate of enterprise innovation and city scale. Statistical Research, 2016, 33(7): 55-63.]
- [33] Kloeckner B R, L'Hostis A, Richard T. Contextual metrics a mathematical definition for a comprehensive approach of geographical distances [J]. Geographical Analysis, 2021, 53(4): 736-766.
- [34] 郭军. 审计师的地理邻近性与客户公司的股价崩盘风险 [D]. 厦门: 厦门大学, 2018. [Guo Jun. The geography of auditors and client firms' stock price crash risk. Xiamen, China: Xiamen University, 2018.]
- [35] 周嘉妮. 分所层面客户重要性、地理因素对内部控制审计质量的影响研究 [D]. 厦门: 厦门大学, 2019. [Zhou Jiani. Office-level client importance, geographic factors and internal control audit quality. Xiamen, China: Xiamen University, 2019.]
- [36] 罗进辉, 李雪, 林芷如. 审计师—客户公司的地理邻近性与会计稳健性 [J]. 管理科学, 2016, 29(6): 145-160. [Luo Jinhui, Li Xue, Lin Zhiru. Effects of geographic proximity between auditors and client firms on accounting conservatism. Journal of Management Science, 2016, 29(6): 145-160.]
- [37] 朱敏, 汪传江. 企业空间区位会影响分析师IPO定价预测的准确性吗? [J]. 财经论丛, 2017(10): 60-70. [Zhu Min, Wang Chuanjiang. Does the location of IPO-firms influence analysts' IPO pricing prediction accuracy? Collected Essays on Finance and Economics, 2017(10): 60-70.]
- [38] 陈青雁, 王鹏, 钟业喜. 中国民营上市企业总部空间格局及影响因素 [J]. 世界地理研究, 2020, 29(5): 996-1005. [Chen Qingyan, Wang Peng, Zhong Yexi. Spatial pattern and influencing factors of headquarters of China's private listed companies. World Regional Studies, 2020, 29(5): 996-1005.]
- [39] 蒋子龙, 王军, 樊杰. 1990—2019年中国上市公司总部分布变迁及影响因素 [J]. 经济地理, 2022, 42(4): 112-121. [Jiang Zilong, Wang Jun, Fan Jie. Distribution change and influencing factors of headquarters of Chinese listed companies from 1990 to 2019. Economic Geography, 2022, 42(4): 112-121.]
- [40] 王晶. 我国宏观经济统计数据质量诊断方法与实证 [J]. 统计与决策, 2018, 34(4): 34-37. [Wang Jing. Diagnosis method and empirical analysis of China's macroeconomic statistic data quality. Statistics & Decision, 2018, 34(4): 34-37.]
- [41] Cortes F, Gomes A, Gopalan R. Corporate inversions and governance [J]. Journal of Financial Intermediation, 2021, 47: 100880. doi: 10.1016/j.jfi.2020.100880.
- [42] 文婧, 张强国, 杜恒, 等. 北京电影产业空间集聚与网络权力分布特征研究 [J]. 地理科学进展, 2019, 38(11): 1747-1758. [Wen Hu, Zhang Qiangguo, Du Heng, et al. Characteristics of spatial agglomeration and network power distribution of Beijing film industry. Progress in Geography, 2019, 38(11): 1747-1758.]
- [43] 汲昌霖. 政治关联是否促进了企业绩效: 基于金融生态环境的视角 [J]. 公司金融研究, 2015(3): 45-58. [Ji Changlin. Was political association support promotion to enterprise performance? A perspective based on the financial ecosystem. Journal of Corporate Finance, 2015(3): 45-58.]

Errors in statistical data from the perspective of corporate location: A case study of A-share listed companies

HU Guojian^{1,2}, LU Yuqi³, ZHONG Yexi^{1,2*}

(1. School of Geography and Environment, Jiangxi Normal University, Nanchang 330022, China;

2. Key Laboratory of Poyang Lake Wetland and Watershed Research, Ministry of Education, Jiangxi Normal University, Nanchang 330022, China; 3. School of Geography Science, Nanjing Normal University, Nanjing 210023, China)

Abstract: Statistical data errors not only mislead the analysis, decision making, and actions of the government, business, and individuals, but also cast doubt on the conclusions of numerous academic studies. Among these data, social-economic data related to businesses is an important target of statistical work, and the determination of business locations is a prerequisite for statistical data on businesses. Currently, the registered address serves as the business location in many social- economic data, but the phenomenon of spatial separation between confirmed business entities and registered addresses may result in significant errors when statistics are conducted based on registered addresses. This article argued that the spatial separation between business entity locations and registered addresses is a new perspective for understanding and studying statistical data errors, providing the possibility to measure actual data errors. In addition to the registered address, listed companies also disclose their office address, which serves as a breakthrough point for quantifying the separation between business entities and registered addresses and studying the statistical errors in data. Taking the office address as the geographical location of A-share listed company headquarters and using various types of data based on this location as the correct values, this study compared and analyzed the errors in data collected based on the registered address (including the number of companies, company indicators, regional indicators, and geographical distances). The study showed that when statistics are generated based on the registered address, the 233 A- share listed companies examined in this study are misclassified in terms of their province (that is, the headquarters are located in a province that is different from the registered address), with most of these companies being located in economically developed regions such as Beijing, Shanghai, and Guangdong, while the distribution of registered addresses is relatively scattered. Influenced by these companies, the average error rate of the number of companies calculated based on the registered address is 14.23% per province, indicating a more imbalanced distribution of the actual headquarter locations of listed companies compared to the registered addresses. The combined average error rate of the six company indicators is 24.84%, showing significant variation in error rates across different provinces and indicators. The average error rate of calculating eight regional development indicators based on the registered address for companies in different industries is 3.04%, indicating a relatively small error rate. The statistical error in calculating the geographical distance between companies and accounting firms based on the registered address depends on the location relationship between the headquarters, registered addresses, and accounting firms. Many companies hire accounting firms in their headquarter locations, leading to a significantly higher geographical distance calculated based on the registered address compared to the actual value. The issue of statistical data errors caused by the spatial separation between business entities and registered addresses will persist in the long term. Given the significant impact of data errors on government decision making, business strategy formulation, academic research, and so on, it is essential for the academic community, especially those in economic geography, to pay attention to the statistical data errors caused by the separation between business entities and registered addresses and their negative effects.

Keywords: enterprise entity; registered address; headquarter; listed company; data error