

# 城市规模分布和经济发展： 存在一种特定模式吗？

陆 旻

**【内容摘要】**基于 Thomas Brinkhoff 的城市人口数据库，采用了 118 个国家相距 10 年的两期数据，检验了国家首位值与经济发展之间的关系。研究发现，由人均 GNI 代表的经济发展水平与首位值之间存在“倒 U 型”关系，即在较低的经济水平上，经济发展与首位值之间的关系为正；而在较高的经济水平上，两者之间的关系为负。同时，经验分析结果也发现了“中心区位理论”在解释城市规模分布中的作用。其中，由经济活动强度和制造业占比代表的规模经济和集聚经济促进人口集中，由人口密度和土地面积代表的土地租金和运输成本促进人口分散。研究结论对中国的启示是，随着中国经济发展水平的提高，特大城市的人口规模并不会无限膨胀，因为“发展”会自动疏散人口。

**【关键词】**位序规模分布；首位分布；经济发展；人口密度

**【作者简介】**陆旻，中国社会科学院人口与劳动经济研究所研究员。北京：100006

## City Size Distribution and Economic Development: Is There a Certain Pattern?

Lu Yang

**Abstract:** We reexamine the relationship between the city size distribution and economic development of a country by using a new data set provided by Thomas Brinkhoff. We have made use of data on 118 countries in two census periods. The empirical results allow us to confirm that there is an inverted-U shape between per capita of income and primacy index. There is a positive relationship between the primacy index and the economic development among the low income countries. However, there is a negative relationship between the two variables among high income countries. This paper also suggests that the central place theory has dominated influence on the city size distribution. Scale and agglomeration economies in production promote urban concentration while land rent and higher transport cost foster population dispersal. The results indicate that population dispersal will accompany with a country's development and the population in primacy city of China will shrink in the future.

**Keywords:** Rank-Size Distribution, Primacy Distribution, Economic Development, Population Density

**Author:** Lu Yang is Professor, Institute of Population and Labor Economics, Chinese Academy of Social Sciences. Email: luyang2002@cass.org.cn

## 1 引言

当一个国家的所有城市按照人口规模进行排序时,将出现不同的城市人口规模分布特征。目前,文献研究中已经识别出两种城市规模分布的类型:第一种城市规模分布类型是 Jefferson(1939)提出的“首位分布”(Primate),即一些小城市受到一个或几个特大城市“支配”,并且中等城市的数量非常有限。我们在现实世界中也会发现,一些国家的特大城市常常是“出类拔萃”的,不仅体现在人口规模上,也体现在全国范围内的影响上。例如,秘鲁第一大城市(首都利马)的人口数量是第二大城市(阿雷基帕)人口数量的近 11 倍。第二种类型是由 Zipf(1949)提出的“位序—规模”(Rank-Size)分布。这种分布呈现的特征是,国家内部城市人口规模(位序)分布符合“截断对数正态”(Truncated Lognormal)分布。例如,目前美国的城市规模分布特征最符合这一标准,即第一大城市的人口是第二大城市人口的近似 2 倍,是第三大城市人口的近似 3 倍,以此类推,城市人口规模是其在该系统中的位序的反函数。

在理论上,Zipf 的假说与 Jefferson 的假说形成了一个“两分法”,即经济欠发达国家与“首位分布”相联系,经济成熟的国家与“对数正态分布”相联系(例如,“位序—规模”分布)。然而,现有的文献似乎并没有明确的证据表明,两者之间存在某种必然的关系。主要原因是城市规模分布的演进过程并不能够由“经济发展”得到全部解释,一个国家的地理特征、历史因素、乃至种族和习俗等一些常常被放入误差项中的变量在这一研究框架中变得异常重要。因此,我们的研究目的就是対经济发展与城市规模分布之间的关系进行再分析。如果控制上述因素后,城市规模分布的确是经济发展的内生结果,那么在理论上将引申出一项令人振奋的政策含义:具有“首位城市”分布特征的国家并不需要采用行政干预的手段改变本国的城市分布格局,因为“发展”将会自动疏散人口。与之前的文献相比,我们的研究在如下两个方面有所推进。

第一,我们采用了 Thomas Brinkhoff 的城市人口数据库<sup>①</sup>。这个数据库中收集和跟踪了 100 多个国家的城市人口数据。因此,我们选择了更多的样本国家和跨越 10 年的两期数据集<sup>②</sup>。更大的样本可以帮助我们分析和验证各种类型的国家城市规模分布(首位值)、变化趋势以及哪些因素影响了这种变化趋势。这与 Berry(1961)等由于受到数据的限制,仅仅采用 38 个国家作为样本的经验分析不同,本文在数据代表性上更有说服力。

第二,我们计算了更完整的首位度指标。虽然 El-Shakhs(1972)提出的城市首位度指标涵盖了一个国家城市系统中的所有城市,但是由于数据限制,El-Shakhs(1972)只是采用了一个简化近似方程计算“首位值”。与之前文献不同,我们采用 El-Shakhs(1972)的完整公式计算了更准确的国家“首位值”。为了与之前的文献做对比,我们还同时计算了“首位城市在前 4 位城市中的占比”指标,并对文献中产生“不理想”结果的原因做出了解释。

虽然很多文献都试图证明人均收入水平与“城市系统”首位值之间的相关性,但是研究结果却往往没有支持理论预期。与之前文献研究结果的最大区别在于:我们发现了由人均 GNI 代表的经济发展水平与首位值之间存在“倒 U 型”关系,而且这种关系是稳定的。同时,

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

① City Population <http://www.citypopulation.de>。

② 国家的城市人口数据基本来自人口普查年份,而逐年数据并不可得。

“中心区位理论”确实可以解释一个国家城市规模分布。其中,由经济活动强度和制造业占比代表的规模经济和集聚经济能够促进人口集中,相反,由人口密度和土地面积所代表的地租和运输成本将促进人口分散。因此,随着中国经济发展水平的提高,特大城市的人口规模并不会无限膨胀,因为“发展”会自动疏散人口。

## 2 文献综述

Zipf(1949)提出了一个著名假说,即城市人口规模分布符合“位序—规模”分布。经过学者的反复验证,发现“位序—规模”分布实际上就是一种特殊形式的“帕累托分布”,即帕累托指数等于1。换句话说,这一假说认为:如果我们将城市从最大到最小规模进行排序,城市的位序和城市人口之间的乘积为常数。

尽管现实中一些国家的城市分布特征确实符合“帕累托分布”,但是很多国家也同样存在另一些分布特征。在世界范围内,人们总能发现一些国家的第一大城市的人口规模是第二大城市人口规模的数倍。Jefferson(1939)将第一大城市人口数量标准化为100,发现第二大城市人口数平均为30,第三大城市人口数平均为20。Jefferson(1939)发现“从世界范围来看似乎在证明这样一个法则:一个国家的首位城市常常是‘过大’,并且在各个方面都对全国其他地区产生重要的影响”,进而Jefferson首次提出了“首位城市”的概念。Mehta(1964)认为Jefferson提出的“首位城市”概念与后来很多学者提到的“首位城市”概念在含义上有所不同。因为在Jefferson看来,像纽约、东京以及其他类似的城市只要在全国范围内的影响上是“出类拔萃”的,就可以称为“首位城市”,即便这个城市的人口规模仅仅是第二大城市人口规模的两倍。然而,在后来很多学者看来,只有当第一大城市是第二大城市人口规模的“数倍”时,才能被称为“首位城市”(Primate Cities)。<sup>①</sup>

基于现实世界中存在的这两种截然不同的城市规模分布类型,我们一定会产生这样的疑问:为什么世界上城市规模分布会存在差异?事实上,对于这种差异的解释存在很多假说。例如,如果影响城市规模分布的重要“力量”比较单一则会形成首位分布,如果影响“力量”很多或者城市发展历史很长则会形成对数正态分布(Berry和Garrison,1958)。此外,还有学者采用“中心区位理论”解释城市规模分布类型,即城市数量和城市规模可以被视为两种相反经济力量的结果:生产效率和分布效率。“由于利润是由生产效率和分布效率所决定,当企业追求利润最大化时,一定会选择一单位的生产成本和一单位的分布成本最小化时的最优水平,此时,城市的均衡数量和均衡规模就形成了;一旦达到这种状态,任何试图向其他城市再选址的企业将会使利润减少”(Alperovich,1993)。这很容易推测出一个“两分法”的世界,即经济欠发达国家与“首位分布”相关,因为影响分布的力量单一且人口向大城市集中效率更高;相反,经济成熟的国家常常与“位序—规模”分布相关,因为影响分布的力量逐渐增加,特别是人口分散更能产生效率。这种观点的另一种表述是,在较低的经济发展水平上,经济发展和首位值之间的关系为正,在较高的经济发展水平上,两者之间的关系为负。以检验“两分法”为目标的文章有Berry(1961)和Alperovich(1993)等。

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

<sup>①</sup> 本文虽然沿用Jefferson(1939)的“首位城市”概念,但是文中提到的“首位分布”都是指Jefferson(1939)之后学者们提到的首位分布含义。

在理论上,虽然经济发展和城市规模分布之间可能存在某种特定的规律,但是经验分析却很少支持这一假说。例如,Berry(1961)采用主成分分析的方法构造了一个衡量经济发展程度的指标,这个指标通过4个纬度的“尺度”对经济发展进行测量。他将其中的“技术尺度”作为横坐标,“人口统计尺度”作为纵坐标,并将城市规模分布的4种类型用散点图的形式描述在“经济—人口发展”坐标中,最终发现,不同类型的城市规模分布在图中并没有展现出某种有规律的“特征”。因此他认为,不同的城市规模分布与国家之间的相对经济发展阶段没有关系,位序分布并不是国家城市体系发展的顶点,首位分布也不是欠发达国家的“专利”。Berry(1961)的研究框架非常具有启发性,但是受限于当时的数据条件,38个国家的样本量还难以得出更准确的结论。Mehta(1964)对人均GNP排序后进行分组并分别计算每组的平均“首位值”,最后并没有发现经济发展和首位值之间存在相关性。虽然研究结论受限于统计描述方法,但Mehta(1964)给出了一个富有启发性的猜想:“首位城市”似乎在很大程度上是“小地区”和“小国”的一个函数。为了使不同国家和同一国家不同时期的城市规模分布具有可比性,El-Shakhs(1972)提出了一个测量城市首位程度的计算公式,根据统一的公式可以计算出相同口径下可比的“首位值”。

事实上,从20世纪80年代开始,文献研究中又出现了一些新的方法,例如,Rosen和Resnick(1980)、Wheaton和Shishido(1981)、Sheppard(1982)、Parr和Jones(1983)、Parr(1985)、Alperovich(1992、1993)和Soo(2005)的研究。这些研究更多偏向于城市规模分布的影响因素分析。文献大多采用“中心区位理论”作为分析框架(Hsu,2012;Düben和Krause,2020),并将影响城市规模分布的“力量”分成4大类:规模经济和集聚经济、市场规模、运输成本和地租。其中,前两个因素将促使人口向“首位城市”集中,而后两个因素使人们远离“首位城市”。但是,已有的经验分析都仅仅支持了某一种或几种因素对首位值的影响,而且预期的影响方向与理论之间并不十分一致。此外,很多经验研究还加入了其他控制变量,El-Shakhs(1972)提到了以下主要的变量:(1)城市化的历史长度以及殖民主义的影响(Berry,1961);(2)领土面积、人口规模以及大国的区域影响(Berry,1961;Mehta,1964);(3)城市化程度以及人口增长率(Browning,1962);(4)人口密度(Linsky,1965);(5)政府干预(Alperovich,1993)。此外,也有文献从其他角度试图对城市规模分布形成的原因进行解释,例如,从城市间的交通网络视角(Berliant和Watanabe,2018)以及从不平等视角(Castells-Quintana,2018)解释城市规模分布。本质上这些研究并没有超出“中心区位理论”的总体框架,即生产效率和空间分散效率之间的长期一般均衡形成了不同的城市规模分布,而经济的集聚效应和分散效应通常与经济发展阶段相关,在理论上经济发展和首位度之间应该存在某种相关性,然而由于数据限制,鲜有文献对两者之间的关系进行检验。

### 3 城市规模分布的测量方法

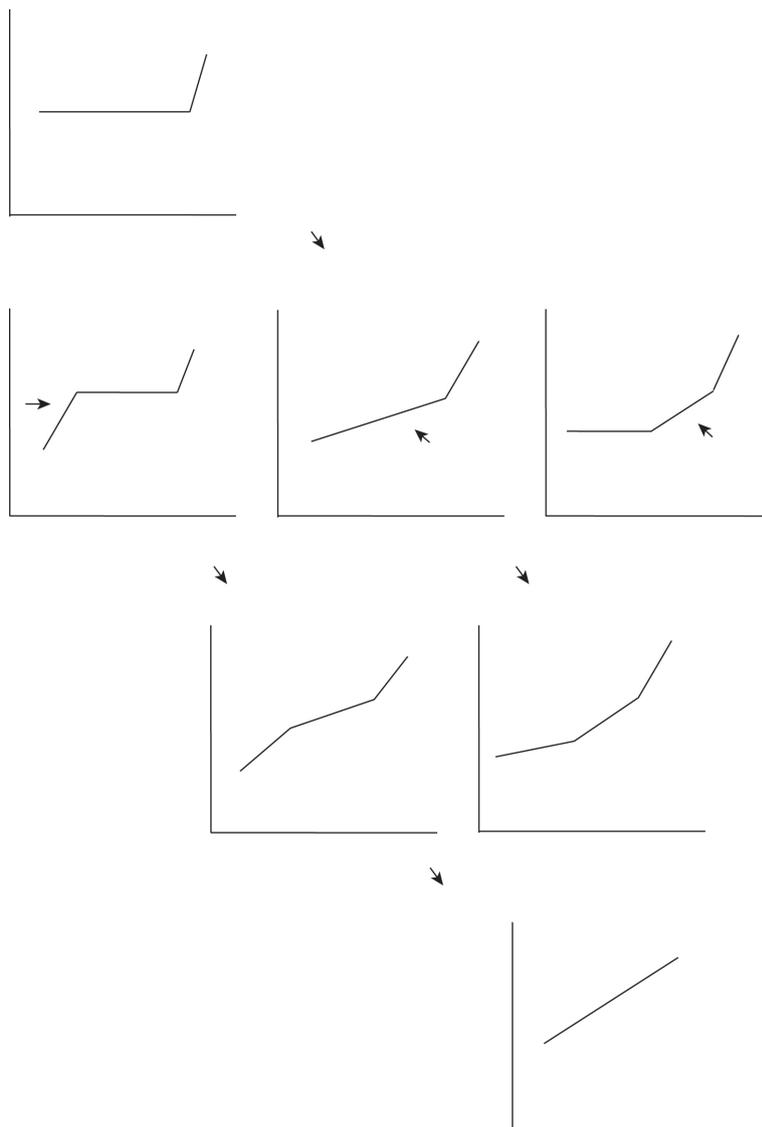
在文献中,对城市规模分布的测量有两种截然不同的方法:分别是Berry(1961)给出的定性描述方法,以及采用公式来衡量国家城市“首位程度”的方法(El-Shakhs,1972)。事实上,每种方法在测量城市规模分布时都不具有完全的优势。

第一种方法是Berry(1961)采用的城市人口累积频数的方法。Berry(1961)用累积频率图表示国家城市人口规模分布类型。基于经验分析,Berry(1961)给出了4种(共7种具体形

态) 城市规模分布类型( 见图 1) 。如果一个城市的规模分布符合对数正态分布, 那么, 其累积频率将呈现为一条直线。从图 1 最顶端( 代表首位程度最高的城市规模分布类型) 向最底端( 代表对数正态分布) 的图形变化过程, 反映了国家城市的首位程度在不断减弱、对数正态分布趋势在不断加强。

图 1 Berry( 1961) 提出的城市规模分布发展模型

Figure 1 The Model of City-Size Distribution and Economic Development in Berry( 1961)



第二种方法是通过测量一个国家城市系统的“首位值”来衡量首位程度。在之前的文献中, 曾经采用过第一大城市相对第二大城市的测量方法( Linsky, 1965)、相对前 3 大城市的测量方法( Jefferson, 1939) 或相对前 4 大城市的测量方法( Berry, 1961; Mehta, 1964)。这种方法的优点在于可以观测短时期内“首位城市”的相对变化率, 但是这些方法存在的共同问题是, 没有考虑到一个国家“城市系统”中的所有城市, 而只是关注了城市人口规模排在前几位的城市。为此, El-Shakhs( 1972) 构建了一个首位度指标。

El-Shakhs(1972) 构建指标的基本思想是,“需要构建一个连续的、对任一系统都可应用的测量方法,这种测量方法将考虑到系统中的每一个城市,而不是单单的几组城市。这种测量方法应该指出,在一个特定系统中,所有城市变动的相对程度。这种相对的规模偏移将适用于任何一个城市系统,并且取值将从系统中最小值(0 或没有偏移)(即所有城市规模相等)到最大值(1 或无限偏移)(即一个假定的系统中,所有的城市人口都集中在一个城市,并且下一个城市的人口规模为零)。任一城市系统的首位程度(或首位值)都将落在这两个极值中间。指标值越大,系统中首位程度越高”。因此,在系统中的一个给定城市  $c_i$  的首位值是  $c_i$  城市与每一个人口规模小于  $c_i$  的城市人口规模的差异的平均率,El-Shakhs(1972) 给出的测量方法如下:

$$p_i = \frac{1}{(n-i)} \left[ \frac{c_i - c_{(i-1)}}{c_i} + \frac{c_i - c_{(i-2)}}{c_i} + \dots + \frac{c_i - c_n}{c_i} \right] \text{ 或}$$

$$p_i = \frac{1}{(n-i)c_i} \sum_{j=(i-1)}^n (c_i - c_j) \quad (1)$$

其中  $c$  代表城市人口规模;  $n$  代表系统中城市数量;  $i$  和  $j$  代表按人口规模降序排列的城市,所以最大城市位序为 1; 最小城市位序为  $n$ 。  $p_i$  代表第  $i$  个城市相对所有人口小于第  $i$  个城市的首位值。整个系统中的首位值由首位指标  $p$  表示,等于所有城市的首位度平均值:

$$p = \frac{1}{(n-1)} \sum_{i=1}^{(n-1)} p_i \quad (2)$$

将  $p_i$  用式(1)替代,则:

$$p = \frac{1}{(n-1)} \sum_{i=1}^{(n-1)} \left[ \frac{1}{(n-i)c_i} \sum_{j=(i-1)}^n (c_i - c_j) \right] \quad (3)$$

El-Shakhs(1972) 认为“应用式(3)这一首位指标时将需要大量的、详细的和可比的数据,这些信息对于很多城市来说都是不可得的,特别是发展中国家。即使可以获得,也只有一些大城市的数据,并且对我们的定义和调查日期来说都不可比”。由于数据的限制,El-Shakhs(1972) 选择了一个简化式。到目前为止,在可获得的数据中,Thomas Brinkhoff 的城市人口数据库收录的数据相对较全,而且拥有每隔 10 年左右的城市人口可比数据,相对以往的研究,我们不仅可以直接采用式(3)作为城市“首位值”的计算公式,而且还能计算各个国家在跨度 10 年间的“首位值”变化。

Jefferson(1939) 使用的首位程度指标都是“一个国家最大城市的人口规模与人口规模前 4 位城市人口总和之比”。但是,这个指标存在一个非常严重的缺陷:这个指标只能比较准确地反映出国家城市系统中只有一个单一首位城市的“首位值”。如果一个国家拥有一个以上的首位城市,那么根据这个指标计算的首位值就会很低。同理,如果仅仅考虑一个“城市系统”中的前几位城市,无论采用哪种方法都无法真实描绘出城市系统的分布情况。以图 1 为例,最顶端的首位程度比第二和第三层图形中的首位程度要更高,但是如果采用上述方法,计算出的结果很可能是相等的,因为这种方法忽略了“中间城市”和“小城市”的实际分布特征。

El-Shakhs(1972) 构建的“首位值”指标对捕捉城市规模分布的“轻微”变化更敏感,而且也能够量化短期内的城市系统变化——城市人口更加集中还是更加分散。为此,本文主要采用

El-Shakhs(1972)的“首位值”指标。此外,为了与之前的文献研究做出对比,我们也选择了“首位城市在前4大城市中占比”作为代理指标进行对比分析。

## 4 模型和数据说明

### 4.1 模型形式和理论依据

在分析城市规模分布的影响因素时,文献中所依托的分析框架大都是“中心区位理论”。当生产集聚在大城市中,规模经济和集聚经济是实现空间生产效率的最重要因素。但与此同时,随着越来越多的经济活动集中于大城市,地租和运输成本增加,企业和居民在大城市中的“好处”将下降。一旦更高的土地价格以及其他的负外部性超过了城市人口集中的好处,企业和居民将会有动机从核心城市向净利益更高的周边地区分散(Alperovich,1993)。这一模型很好地说明了基本的权衡关系:“集中”所产生的生产效率和“空间分散”所产生的运输节约。从这一理论框架中得到的基本假说是:任何一个国家的最终城市规模和城市数量可以由4大类因素得到解释:(1)市场规模;(2)规模经济和集聚经济;(3)运输成本;(4)地租。其中,第(1)和第(2)类因素描述了:生产集中可以获利;第(3)和第(4)类因素说明了成本增加导致生产的空间分散。我们可以将模型设置为如下形式:

$$Primacy_i = \alpha + \beta_1 PGNI_i + \beta_2 PGNI_i^2 + \beta_3 Scale_i + \beta_4 IND_i + \beta_5 DEN_i + \beta_6 Land_i + \beta_7 X_i + \varepsilon_i \quad (4)$$

其中, $i$ 代表国家; $Primacy$ 代表首位值; $PGNI$ 代表人均收入水平; $Scale$ 代表国家的市场规模; $IND$ 代表国家的规模经济和集聚经济水平; $DEN$ 代表地租成本; $Land$ 代表国家的运输成本; $X$ 是其他控制变量,包括城市化水平和政府干预; $\varepsilon_i$ 为误差项。在此基础上我们将对以下两种假说进行检验:

假说1:经济发展水平与城市系统的首位程度之间存在两个不同的阶段。在较低的发展水平上,经济发展与首位值之间的关系为正,而在较高的发展水平上,两者之间的关系为负。

假说2:土地面积越大的国家,首位城市对其他城市的经济辐射作用和影响力就越弱,其首位值就会越低;相反,“小国”更倾向具有“首位城市”结构。

### 4.2 样本选择和数据说明

文章中所有城市人口数据均来自 Thomas Brinkhoff 的城市人口数据库;其他相关数据来自世界银行发展指标数据库(World Development Indicators, WDI)。我们选择2000年和2010年两次普查年份的人口数据。由于数据缺失,在5年之内的数据都被认为是有效的,其他解释变量都根据城市首位值的年份进行了相应调整,全部的样本国家有118个。代理变量的选取和描述性统计结果见表1。

首位值作为被解释变量主要基于 El-Shakhs(1972)公式计算得出,并将“首位城市在前4位城市的占比”指标作为辅助代理变量。人均收入水平作为主要兴趣变量,分别采用“人均收入的对数”(美元,2000年不变价格)作为连续变量,以及世界银行定性分组作为虚拟变量<sup>①</sup>。此外,已有文献中通常采用GDP总量代表一个国家的市场规模,规模经济的预期符号为“正”。制造业占比更高的国家其集聚经济也会越大,我们将其作为集聚经济的代理变量,

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

① 世界银行在每个财政年度(通常是每年的7月1日到第二年的6月30日)对所有国家的收入分组进行调整,数据来自 <http://data.worldbank.org/about/country-classifications/a-short-history>。

预期符号为“正”。采用其他文献中的做法 我们将人口密度变量作为地租的代理变量 其预期符号为“负”。我们采用土地面积作为运输成本的代理变量。首先,土地面积是最具外生性的因素,土地面积越小的国家,其城市系统中的“首位城市”对其他城市的经济辐射作用和影响力就越强,形成首位分布的可能性就越大;其次,土地面积越小的国家往往运输距离会越短,在相同的条件下,运输成本就会越低,人口“空间分散”效应就越低,运输成本的预期符号为“负”。

表 1 变量定义及描述性统计

Table 1 Definition and Description of Variables

变量名称	代码	变量描述	数量	平均值	中位值	最大值	最小值
首位值	<i>Primacy_Total</i>	根据公式(3)计算的首位值 $\times 100$	216	39.42	38.16	64.94	24.42
	<i>Primacy_Four</i>	最大城市人口/前4位城市人口(%)	230	55.91	55.89	87.68	25.93
人均收入 (连续值)	<i>PGNI</i>	人均国民收入水平(美元)	204	8684	2716	53507	78
	$\log(PGNI)$	人均国民收入水平的自然对数形式	204	8.01	7.91	10.89	4.36
收入水平 (哑变量)	<i>L</i>	低收入国家 $L = 1$ , 否则 $L = 0$	200	0.19	0	1	0
	<i>LM</i>	中低收入国家 $LM = 1$ , 否则 $LM = 0$	200	0.28	0	1	0
	<i>UM</i>	中高收入国家 $UM = 1$ , 否则 $UM = 0$	200	0.21	0	1	0
	<i>H</i>	高收入水平国家 $H = 1$ , 否则 $H = 0$	200	0.32	0	1	0
市场规模	<i>GDP</i>	国内生产总值(千亿美元)	222	3.14	0.34	115.98	0.006
	<i>GDP/Land</i>	(千亿美元/万平方公里)	222	0.29	0.013	24.65	7.3E-05
集聚经济	<i>IND</i>	制造业增加值/国内生产总值	209	0.16	0.16	0.38	0.03
地租	<i>DEN</i>	人口/土地面积(人/1000平方米)	230	0.17	0.07	7.25	0.002
运输成本	<i>Land</i>	土地面积(10平方公里)	236	0.096	0.022	1.64	3.2E-05
城市化率	<i>URB<sub>1</sub></i>	城市人口/总人口	230	0.60	0.61	1	0.09
	<i>URB<sub>2</sub></i>	城市化率 $> 50\%$ $URB_2 = 1$ , 否则 $URB_2 = 0$	230	0.71	1	1	0
政府干预	<i>GOV</i>	政府消费/国内生产总值	217	0.16	0.17	0.34	0.04
亚洲	<i>ASIA</i>	属于亚洲 $ASIA = 1$ , 否则 $ASIA = 0$	236	0.30	0	1	0

续表1

变量名称	代码	变量描述	数量	平均值	中位值	最大值	最小值
欧洲	<i>EURO</i>	属于欧洲 $EURO = 1$ ,否则 $EURO = 0$	236	0.30	0	1	0
美洲	<i>AMER</i>	属于美洲 $AMER = 1$ ,否则 $AMER = 0$	236	0.19	0	1	0
非洲	<i>AFRI</i>	属于非洲 $AFRI = 1$ ,否则 $AFRI = 0$	236	0.17	0	1	0

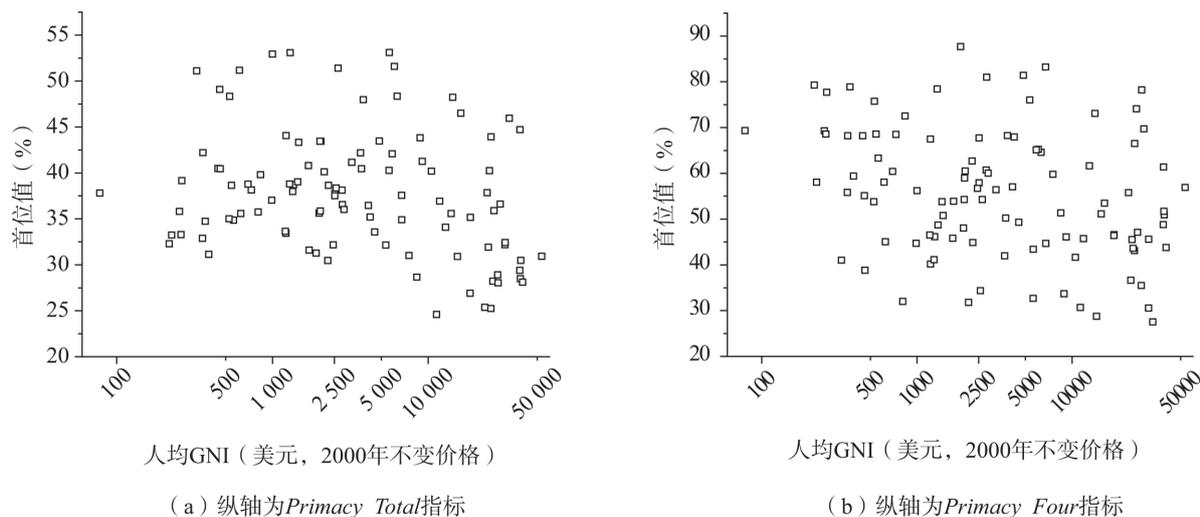
## 5 经验分析结果

### 5.1 城市首位程度的其他影响因素

图 2 给出了 2010 年人均收入与首位值之间的关系(纵轴代表首位值,横轴代表人均收入,横轴刻度采用了对数刻度)。从散点图看出,基于 El-Shakhs(1972)公式计算的首位值与人均收入之间存在倒 U 型的关系;传统首位值指标与人均收入之间呈现出负相关关系。

图 2 2010 年 103 个国家的人均 GNI 与城市首位值之间的关系

Figure 2 The Relationship between GNI Per Capita and Primacy of 103 Countries in 2010



人均GNI (美元, 2000年不变价格)

(a) 纵轴为Primacy\_Total指标

人均GNI (美元, 2000年不变价格)

(b) 纵轴为Primacy\_Four指标

基本模型结果显示,当 *Primacy\_Total* 作为被解释变量时,人均收入的一次项显著为正,二次项显著为负;当 *Primacy\_Four* 作为被解释变量时,人均收入一次项显著为负,二次项不显著<sup>①</sup>。说明采用前 4 位城市人口数据构造的首位值并不能反映城市系统全貌。因此,估计结果在很大程度上仅度量了经济发展对第一大城市人口相对规模的影响。但在很多情况下,第一大城市人口相对减少,可能伴随了第二大城市或第三大城市人口规模的相对增加。对整体城市系统来说,城市体系的偏离程度并没有改善。这也就解释了之前文献采用简化公式难以发现首位值与经济发展水平之间存在相关关系的原因。

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

① 由于篇幅限制这里不再列出基本模型回归结果,如需要可与作者联系。

表 2 城市系统首位度的影响因素分析

Table 2 The Determination Factors of City-Size Distribution

解释变量	被解释变量为 <i>Primacy_Total</i>				被解释变量为 <i>Primacy_Four</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
对数人均收入	6.37** (3.09)	7.47** (3.59)	6.51** (3.06)	7.75*** (2.97)	-3.48*** (0.66)	-2.49** (1.04)	-3.33*** (0.69)	-3.19*** (0.71)
对数人均收入平方	-0.47** (0.19)	-0.52** (0.21)	-0.47** (0.19)	-0.54*** (0.19)	—	—	—	—
市场规模	2.04** (0.99)	2.05** (0.99)	2.11** (1.01)	2.37** (1.07)	3.92** (1.72)	3.91** (1.66)	3.96** (1.74)	4.32** (1.77)
集聚经济	13.17* (7.91)	12.45 <sup>a</sup> (7.88)	13.32 <sup>a</sup> (8.19)	14.69* (7.97)	-4.52 (14.60)	-5.15 (14.25)	-4.21 (15.30)	-0.29 (15.58)
地租	-4.99 <sup>a</sup> (3.41)	-4.91 <sup>a</sup> (3.39)	-5.37 <sup>a</sup> (3.46)	-6.18* (3.69)	-14.98*** (5.68)	-14.67*** (5.53)	-15.22*** (5.75)	-16.40*** (5.90)
运输成本	-2.60* (1.33)	-2.48* (1.34)	-2.62* (1.34)	-1.93 <sup>a</sup> (1.25)	-9.50*** (3.55)	-9.10** (3.82)	-9.57*** (3.60)	-8.63** (3.60)
城市化率	—	-2.91 (3.99)	—	—	—	-9.59 (8.08)	—	—
政府干预	—	—	-13.53 <sup>a</sup> (8.47)	-80.15* (40.73)	—	—	-9.47 (16.49)	-114.31 <sup>a</sup> (78.19)
政府干预平方	—	—	—	187.02 <sup>a</sup> (116.5)	—	—	—	298.83 (215.31)
亚洲	-8.93*** (1.98)	-8.99*** (1.97)	-8.97*** (2.07)	-9.64*** (2.15)	9.64** (4.08)	9.91*** (3.70)	9.39** (4.18)	8.44** (4.20)
欧洲	-11.8*** (1.91)	-11.8*** (1.89)	-11.46*** (1.97)	-11.69*** (2.08)	15.42*** (4.02)	15.71*** (3.66)	15.44*** (4.13)	15.04*** (4.08)
美洲	-6.00*** (1.97)	-5.76*** (1.99)	-6.11*** (2.02)	-7.11*** (2.15)	13.05*** (4.30)	14.49*** (3.97)	12.79*** (4.40)	11.59*** (4.40)
非洲	-8.50*** (2.07)	-8.61*** (2.06)	-8.26*** (2.12)	-8.27*** (2.23)	8.23** (4.17)	8.16** (3.73)	8.18* (4.26)	8.03* (4.24)
常数项	26.39** (11.89)	22.77* (13.24)	27.11** (11.78)	27.41** (11.57)	74.85*** (6.59)	72.23*** (6.77)	75.27*** (6.92)	82.25*** (8.04)
R <sup>2</sup>	0.316	0.318	0.330	0.342	0.232	0.238	0.231	0.240
观测值	186	186	185	185	198	198	197	197

注: ①括号中的数据为标准差; ②\*\*\* \*\* \* 分别表示在 1% 5% ,10%水平上显著; ③标记“a”符号的变量代表在 15%水平上显著,虽然在通常情况下超过 10%水平上显著的变量被视为没有通过显著性检验,但是因为显著水平非常接近 10%,为了与其他变量有所区别,在此做出标记。

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net  
在此基础上,我们在模型中加入“中心区位理论”的 4 个主要变量(见表 2)。*Primacy\_Total* 作为被解释变量时,人均收入的一次项和二次项仍然显著。市场规模变量和集聚经济变量分别在 5%和 10%的显著水平下通过了检验,而且这两个变量对首位值的影响为正向,

与我们的理论预期相同。由人口密度作为“地租”的代理变量时,符号为“负”,与理论预期相同。显著性水平虽超过了 10%,但低于 15%。当土地面积作为运输成本代理变量时,变量在模型中显著为负,同样与理论预期相符。此外,无论是连续变量还是虚拟变量,城市化率在模型中都不显著。同样,政府干预变量在模型中也并不显著。*Primacy\_Four* 作为被解释变量时,人均收入和首位值之间存在显著的负相关关系,其他变量估计结果与 *Primacy\_Total* 作为被解释变量时相似。

## 5.2 稳健性检验

为了进一步对上述主要结论的稳健性进行验证,我们将回归方程做出如下调整。首先,将人均收入的连续变量替换成虚拟变量,世界银行每个财政年度会针对不同类型国家提供不同“援助”,并将所有国家的人均 GNI 根据 *Atlas Methodology* 标准进行分组。因此,我们对人均收入组的划分标准采用了世界银行的标准,即每年 7 月 1 日对所有成员国以及人口在 3 万人以上的经济体进行的经济发展定性分组。加入虚拟变量可以区分不同国家所处的发展阶段。根据这一标准,我们将世界上所有国家分成 4 类:低收入国家(L)、中低收入国家(LM)、中高收入国家(UM)、高收入国家(H),其中,中高收入组(UM)作为参照组。模型设定如下。

$$Primacy_i = \alpha + \beta_1 L_i + \beta_2 LM_i + \beta_3 H_i + \beta_4 Scale_i + \beta_5 IND_i + \beta_6 DEN_i + \beta_7 Land_i + \beta_8 X_i + \varepsilon_i \quad (5)$$

从表 3 估计结果发现,高收入国家(H)的首位值(*Primacy\_Total* 指标)显著低于中高收入国家的首位值,但是低收入国家和中低收入国家的首位值与中高收入国家相比没有显著差异。这可能与采用了人均收入虚拟变量的方法有关,因为世界银行的人均收入划分标准可能与“首位值—人均收入”的变化阶段不相符。“中心区位理论”中的 4 个解释变量与表 2 的估计结果非常相似:市场规模和集聚经济对首位值产生显著正影响;运输成本对首位值产生显著负影响,而地租和政府干预变量在模型中并不显著。*Primacy\_Four* 作为被解释变量时,表 3 与表 2 的结论相同:低收入国家的首位值显著高于中高收入国家,高收入国家的首位值显著低于中高收入国家,但是中高收入国家与中低收入国家的首位值没有统计上的显著差异。“中心区位理论”的 4 个代理变量的显著性和符号也与表 2 相同,同样由制造业占比代表的集聚经济在模型中并不显著。此外,政府干预对首位度也没有产生预期的显著影响。

上述分析中,对于大多数样本国家来说数据都出现了两次,分别在 2000 年和 2010 年,当然有些国家因为数据可得性问题样本年份有一定偏差。为了检验结论的稳健性,我们将两期数据分别做截面回归,从而观测主要结论是否与样本的选择时期有关。表 4 显示,以 2010 年多国数据作为样本时,变量的显著性和影响方向与之前的分析一致。当人均收入采用连续变量时,人均收入与首位值(*Primacy\_Total*)之间仍然呈现了稳定的倒 U 型关系。但是,集聚经济变得不再显著;由土地面积代表的运输成本变量显著性水平略超过了 10%,不过仍然在 15% 的显著性水平下通过检验;在任何一种模型形式下,政府干预变量都不显著。如果采用虚拟变量代表收入水平时,高收入国家的首位程度依然显著低于中高收入组;市场规模和集聚经济都显著促进人口集中,相反,地租和运输成本显著促使人口分散,进而支持了我们前面的基本结论。但是,如果我们将 2000 年时间点上的多国截面数据作为样本,除了人均收入与 *Primacy\_Total* 之间仍然存在倒 U 型关系之外(特别是采用人均收入虚拟变量后也同样支持了这

样的结论,“中心区位理论”中的 4 个主要变量在模型中都不显著。我们认为,这可能是由以下两个原因导致的结果:第一,有效观测值减少,与 2010 年样本相比,2000 年样本的实际观测值平均减少了 15 个;第二,在 2000 年样本中,样本国家的数据采集时点更加分散。但是,各主要变量对城市首位值的影响方向都与理论预期相符。此外,我们将样本国家按照人口规模进行分组,主要结论均未发生改变。

表 3 稳健性检验: 经济发展虚拟变量

Table 3 Robustness Test: Dummy Variables of Economic Development

解释变量	被解释变量为 <i>Primacy_Total</i>				被解释变量为 <i>Primacy_Four</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
低收入国家	-0.55 (1.43)	0.15 (1.44)	-0.25 (1.41)	-0.60 (1.38)	14.98*** (3.25)	13.94*** (3.29)	13.43*** (3.29)	12.64*** (3.35)
中低收入国家	0.54 (1.22)	0.43 (1.23)	-0.09 (1.19)	-0.35 (1.19)	2.84 (2.73)	1.85 (2.72)	1.35 (2.76)	0.93 (2.73)
高收入国家	-3.13** (1.48)	-3.72** (1.57)	-3.39** (1.61)	-3.48** (1.62)	-4.46* (2.49)	-4.06* (2.43)	-3.76 <sup>a</sup> (2.52)	-3.94 <sup>a</sup> (2.47)
市场规模	—	1.69* (0.97)	1.79* (0.98)	2.00* (1.03)	—	3.86** (1.64)	3.95** (1.66)	4.32** (1.74)
集聚经济	—	13.91* (7.68)	15.45** (7.62)	17.37** (7.42)	—	3.05 (14.81)	4.57 (15.35)	8.01 (15.62)
地租	—	-4.23 (3.38)	-4.69 (3.43)	-5.38 <sup>a</sup> (3.62)	—	-15.19*** (5.53)	-15.62*** (5.62)	-16.80*** (5.93)
运输成本	—	-2.88* (1.47)	-2.92* (1.51)	-2.37 <sup>a</sup> (1.48)	—	-8.96** (3.63)	-9.05** (3.71)	-8.14** (3.69)
政府干预	—	—	-13.66 <sup>a</sup> (8.85)	-69.75* (39.76)	—	—	-14.03 (17.67)	-113.94 (81.19)
政府干预平方	—	—	—	157.74 (111.23)	—	—	—	283.57 (223.30)
亚洲	-7.84*** (2.66)	-9.33*** (2.31)	-9.34*** (2.39)	-9.92*** (2.47)	4.95 (4.36)	6.98 <sup>a</sup> (4.81)	6.64 (5.01)	5.76 (5.00)
欧洲	-10.95*** (2.56)	-12.41*** (2.13)	-12.14*** (2.18)	-12.40*** (2.28)	14.79*** (4.09)	13.16*** (4.68)	13.15*** (4.93)	12.79*** (4.87)
美洲	-5.63** (2.72)	-6.87*** (2.32)	-6.83*** (2.36)	-7.63*** (2.49)	11.43** (4.60)	11.33** (4.89)	11.07** (5.08)	9.83* (5.05)
非洲	-7.57*** (2.76)	-8.77*** (2.36)	-8.59*** (2.40)	-8.73*** (2.50)	5.36 (4.66)	5.06 (4.98)	4.95 (5.17)	4.84 (5.14)
常数项	47.70*** (2.76)	47.33*** (2.66)	49.37*** (3.06)	54.04*** (4.51)	43.64*** (4.47)	45.93*** (5.27)	48.37*** (6.21)	56.49*** (8.85)
R <sup>2</sup>	0.238	0.283	0.302	0.311	0.194	0.260	0.263	0.270
观测值	190	184	183	183	200	194	193	193

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

注: ①括号中的数据为标准差; ②\*\*\* \*\* \* 分别表示在 1%、5%、10%水平上显著; ③标记“a”符号的变量代表在 15%水平上显著,虽然在通常情况下超过 10%水平上显著的变量被视为没有通过显著性检验,但是因为显著水平非常接近 10%,为了与其他变量有所区别,在此做出标记。

表 4 稳健性检验: 被解释变量为首位指标( *Primacy\_Total* )  
Table 4 Robustness Test: *Primacy\_Total* as the Dependent Variable

解释变量	2010 年截面数据				2000 年截面数据			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
人均收入对数	7.66 <sup>*</sup> (4.00)	7.89 <sup>*</sup> (4.03)	—	—	6.77 <sup>a</sup> (4.64)	7.24 <sup>a</sup> (4.43)	—	—
人均收入对数平方	-0.56 <sup>**</sup> (0.25)	-0.57 <sup>**</sup> (0.25)	—	—	-0.50 <sup>*</sup> (0.29)	-0.51 <sup>*</sup> (0.28)	—	—
收入虚拟变量	参照组: 中高收入国家				参照组: 中低收入组			
低收入国家	—	—	-0.11 (2.14)	-0.10 (2.14)	—	—	-2.64 <sup>a</sup> (1.71)	-2.43 <sup>a</sup> (1.63)
中低/中高收入国家	—	—	-1.49 (1.62)	-1.48 (1.63)	—	—	-2.85 (2.12)	-1.59 (2.15)
高收入国家	—	—	-4.93 <sup>**</sup> (2.05)	-4.41 <sup>**</sup> (2.15)	—	—	-5.92 <sup>***</sup> (1.87)	-4.65 <sup>**</sup> (2.02)
市场规模	4.13 <sup>***</sup> (0.72)	4.33 <sup>***</sup> (0.72)	3.31 <sup>***</sup> (0.59)	3.54 <sup>***</sup> (0.66)	2.73 (2.80)	2.78 (2.64)	1.88 (3.19)	2.12 (3.06)
集聚经济	16.76 (12.29)	13.91 (13.05)	23.09 <sup>**</sup> (11.39)	20.76 <sup>*</sup> (11.87)	5.59 (11.64)	9.12 (12.18)	1.89 (11.61)	6.99 (11.92)
地租	-12.67 <sup>***</sup> (2.42)	-13.46 <sup>***</sup> (2.43)	-10.30 <sup>***</sup> (2.06)	-11.21 <sup>***</sup> (2.33)	-5.25 (6.73)	-5.72 (6.37)	-3.36 (7.68)	-4.26 (7.41)
运输成本	-2.69 <sup>a</sup> (1.62)	-2.58 <sup>a</sup> (1.58)	-3.40 <sup>*</sup> (1.84)	-3.27 <sup>*</sup> (1.82)	-2.80 (2.16)	-2.89 (2.27)	-3.20 <sup>a</sup> (2.08)	-3.26 (2.32)
政府干预	—	-12.73 (11.68)	—	-11.78 (12.32)	—	-20.62 <sup>a</sup> (13.39)	—	-18.73 (15.14)
常数项	21.91 <sup>a</sup> (14.93)	22.80 <sup>a</sup> (14.79)	47.96 <sup>***</sup> (3.22)	49.90 <sup>***</sup> (4.02)	26.44 <sup>a</sup> (17.93)	25.54 <sup>a</sup> (16.97)	51.53 <sup>***</sup> (4.81)	52.74 <sup>***</sup> (4.97)
大洲虚拟变量	有	有	有	有	有	有	有	有
R <sup>2</sup>	0.327	0.334	0.294	0.301	0.384	0.417	0.374	0.405
观测值	100	100	100	100	86	85	84	83

注: ①括号中的数值为标准差; ②\*\*\* \*\* \* 分别表示在 1%、5%、10%水平上显著; ③标记“a”符号的变量代表在 15%水平上显著, 虽然在通常情况下超过 10%水平上显著的变量被视为没有通过显著性检验, 但是因为显著水平非常接近 10%, 为了与其他变量有所区别, 在此做出标记。

## 6 结论和不足

本文检验了城市规模分布与经济发展之间的统计关系是否符合发展经济学“两分法”的观点, 而之前的大多数经验研究没有发现任何证据支持这一假说。为此, 我们根据 El-Shakhs (1972) 构建的首位度指标, 采用完整公式计算了 118 个国家在 2000 年和 2010 年的城市系统

首位值<sup>①</sup>, 并得到如下结论: Electronic Publishing House. All rights reserved. <http://www.cnki.net>

① 为节省篇幅, 不再列出各国首位值数据, 如需要可与作者联系。

第一,与以往研究结果的最大区别在于:我们发现了人均收入与首位值之间存在“倒 U 型”关系,进而支持了发展经济学中的“两分法”观点,即在较低的发展水平上,经济发展与首位值之间的关系为正,而在较高的发展水平上,两者之间的关系为负。

第二,由经济活动强度和制造业占比代表的规模经济和集聚经济能够促进人口集中,而人口密度和土地面积所代表的地租和运输成本将促使人口分散。这 4 个影响因素作为控制变量加入模型,结果与预期影响方向相同。但是,预期对城市规模分布产生影响的城市化率和政府干预变量在模型中并不显著。

需要说明的是,本文在方法论上更接近于库兹涅茨曲线假说和环境库兹涅茨曲线假说,是对经济发展与特殊变量之间的一个统计现象的描绘,因此难以进行因果关系识别。然而与上述两个假说不同的是,由于数据的可得性问题,即使在统计描述的层面,城市规模分布和经济发展之间的“倒 U 型”关系也缺乏数据支持。本文的研究出发点和主要贡献在于采用更全面系统的城市人口数据计算了城市偏移度,从统计现象中发现了经济发展水平和城市首位值之间确实存在着“倒 U 型”关系。

如果城市规模分布和经济发展之间的“倒 U 型”关系得到更多数据支持,那么其带来的政策启示也十分明显——经济发展会自动疏散人口。随着中国从中高收入国家进入高收入国家,城市规模分布的首位值会降低,人口也会从特大城市逐渐流入其他中小城市,因为经济发展本身具有人口分散的内在动力。值得注意的是,在经验分析部分,本文也同样存在一些不足:首先,城市人口数据库中均为规模以上城市,因此并没有包括一些人口规模很小的城市。数据更加偏向规模以上城市的变化,因此,如果大城市人口向外分散时更多地流入到这些小城市,那么指标将不能识别这种变化。其次,我们采用了制造业增加值占比指标作为集聚经济的代理变量,虽然这是文献中通常采用的方法,但由此也会引起指标代表性问题。再次,针对某一国家进行长期观测也十分重要,例如,观测中国城市规模分布类型变化或者首位值的变化,这也是未来可以进一步研究的方向。

最后,我们还想强调的问题是,“首位分布”与“对数正态分布”都是城市分布的一种形态。在很多情况下,“首位分布”是否对一国经济产生不利的影响?这恐怕还要取决于经济发展阶段以及短期和长期效应。正如 Mehta(1964)所说“即使特大城市可能是‘寄生的’,但大城市的优势不能被忽视。当然,在此过程中涉及的大量经济和社会成本以及财富是存在浪费的,但是并不能确定‘首位城市’在所有时期的平衡表中总显示为‘红色’。进一步讲,即使我们承认在目前的情况下,这些城市可能是‘寄生的’,但是从长期来看,这些特大城市对本国社会和经济发展的影响是‘有生产力的’。”

---

#### 参考文献/References:

- 1 Alperovich G. 1992. Economic Development and Population Concentration. *Economic Development and Cultural Change* 1: 63-74. (C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>
- 2 Alperovich G. 1993. An Explanatory Model of City-size Distribution: Evidence from Cross-Country Data. *Urban Studies* 9: 1591-1601.

- 3 Berry B. 1961. City Size Distributions and Economic Development. *Economic Development and Cultural Change* 4: 573–588.
- 4 Berry B. and Garrison W.L. 1958. The Functional Bases of the Central Place Hierarchy. *Economic Geography* 2: 145–154.
- 5 Browning C.E. 1962. *Urban Systems and Economic Development*. Eugene: University of Oregon Press: 16–27.
- 6 Berliant M. and Watanabe A.H. 2018. A Scale-Free Transportation Network Explains the City-Size Distribution. *Quantitative Economics* 3: 1419–1451.
- 7 Castells-Quintana D. 2018. Beyond Kuznets: Inequality and the Size and Distribution of Cities. *Journal of Regional Science* 3: 564–580.
- 8 Düben C. and Krause M. 2020. Population, Light, and the Size Distribution of Cities. *Journal of Regional Science* 1: 189–211.
- 9 El-Shakhs S. 1972. Development, Primacy, and Systems of Cities. *The Journal of Developing Areas* 1: 11–36.
- 10 Hsu Wen-Tai. 2012. Central Place Theory and City Size Distribution. *Economic Journal* 563: 903–932.
- 11 Jefferson M. 1939. The Law of the Primate City. *Geographical Review* 2: 226–232.
- 12 Linsky A.S. 1965. Some Generalizations Concerning Primate Cities. *Annals of the Association of American Geographers* 3: 506–513.
- 13 Mehta S.K. 1964. Some Demographic and Economic Correlations of Primate Cities: A Case of Reevaluation. *Demography* 1: 136–147.
- 14 Parr J.B. 1985. A Note on the Size Distribution of Cities Over Time. *Journal of Urban Economics* 2: 199–212.
- 15 Parr J.B. and Jones C. 1983. City Size Distributions and Urban Density Functions: Some Interrelationships. *Journal of Regional Science* 3: 283–307.
- 16 Rosen K.T. and Resnick M. 1980. The Size Distribution of Cities: An Examination of the Pareto Law and Primacy. *Journal of Urban Economics* 2: 165–186.
- 17 Sheppard E. 1982. City Size Distributions and Spatial Economic Change. *International Regional Science Review* 2: 127–151.
- 18 Soo K.T. 2005. Zipf's Law for Cities: A Cross-Country Investigation. *Regional Science and Urban Economics* 3: 239–263.
- 19 Wheaton W.C. and Shishido H. 1981. Urban Concentration, Agglomeration Economies, and the Level of Economic Development. *Economic Development and Cultural Change* 1: 17–30.
- 20 Zipf G. K. 1949. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge: Addison-Wesley.

(责任编辑: 宋 严 收稿时间: 2021-03)