

# 经济分布基准下的中国人口 分布均衡测度研究<sup>\*</sup>

——基于 Matlab 空间模拟的估算

尹虹潘 刘渝琳 刘姝伶

**【摘 要】**文章提出了经济分布基准下的人口分布均衡度测度指标,并依据相关经济理论设计了基于 Matlab 的空间模拟与估算方法。选取 89 个地级以上行政单元作为节点,对中国 2005~2012 年的人口与经济分布进行空间模拟,反推得到的总人口和国内生产总值与统计值总体吻合。以此为基础的估算结果表明,2005~2012 年全国以经济分布为基准的人口分布均衡度总体上在不断提高,同期人口与经济重心距离趋于缩小佐证了这一结论。同时,部分参数的变化从侧面反映出研究期内全国的人口与经济都处于集聚化发展趋势中,因此建议进一步扫除人口与其他要素的流动阻碍,通过集聚来实现人口与经济在更高发展水平上的分布均衡。

**【关键词】**人口分布 经济分布 均衡度 空间模拟

**【作 者】**尹虹潘 重庆大学经济与工商管理学院,副研究员;刘渝琳 重庆大学公共管理学院,教授;刘姝伶 重庆市渝北区统计局,统计师。

## 一、引 言

随着近年来中国经济转型、劳动力供求关系变化(特别是第六次全国人口普查后),经济学界和人口学界都更加关注人口与经济的长期均衡发展,并提出了建设人口均衡型社会的理念,其中包括规模、结构等多个方面,而人口分布均衡也是人口与经济实现长期均衡发展所关注的重点之一。一般意义上,人口分布均衡可以理解为人口分布与地理空间或其他因素(如生态环境承载容量、经济发展等)分布的匹配情况。当前,中国正处于新型城镇化发展的重要时期,而作为幅员辽阔的大国,国内地区差距客观上也将长期存在,经济因素成为

<sup>\*</sup> 本文为重庆市社会科学规划项目“新型城镇化进程中的人口有序流动及其分布均衡研究”(项目编号:2013PYLJ03)的阶段性成果。

影响人口乡城与区域流动及人口分布演变最重要的微观动力之一,这是由广大人民都希望过上更加富足的生活决定的。鉴于此,本文探讨的人口分布均衡并不是人口在地理空间上的均匀分布或是人口与其他因素分布情况的匹配,而是特指人口分布与经济分布的空间匹配,即以同期经济分布情况为基准来衡量的人口分布均衡程度,因此本文所分析的人口分布均衡也隐含了区域经济发展的均衡问题。

我们希望知道,以经济分布情况为基准来看中国当前的人口分布是否均衡,它呈现出怎样的演变趋势,应该如何来测度这种均衡水平?对此,近年来出现了较多的相关研究。其中,基于区域人口份额与经济份额偏离情况分析二者不均衡的程度(张车伟、蔡翼飞,2013;封志明、刘晓娜,2013)是一种常见的方法,可较好地对比不同区域人口与经济两方面因素。用全国人口重心、经济重心的位置变化轨迹来判断人口分布与经济分布是否趋于协调(郭斌、文雯,2013;胡安俊、刘元春,2013)也是现有研究中采用较多的方法,并且这种方法真正考虑了空间因素,相对纯粹的区域份额分析也有其合理之处。

然而,已有文献大多以省级行政区域作为基本单元,这样的做法虽然是受限于数据可得性,但由于中国省级行政区域地理尺度差别太大,研究结果往往准确度不够高。此外,这些研究方法还存在各自的不足。比如,人口与经济份额偏离分析法没有充分考虑空间因素,并且主要是揭示人口与经济分布的不均衡程度(而不是均衡程度);单纯的重心分析法不能完全满足对于人口分布均衡问题的研究需要。首先,重心分析的结论只能回答人口分布均衡程度可能的变化方向(是提高还是降低),却无法回答均衡程度的具体水平和变化幅度。其次,重心分析没纳入分布细节信息,有时会产生误导。一个极端例子是区域内的人口全部均匀地分布在地理空间上,而经济全部集中于人口重心上,按照重心分析来看此时的人口与经济重心是完全重叠的,如果以此认为人口分布相对于经济分布是均衡的就错了。现实中不太会有这种极端情况,但至少表明单纯根据重心分析做出判断并不完全可靠。于是,我们希望用细化的人口分布和经济分布来进行分析,这样应该比基于省级行政区域数据的研究更加准确,并且细化数据可以克服难以考虑空间细节因素引起的可靠性问题。然而,这样又产生了两个新的问题:一是用什么样的指标来衡量经济分布基准下的人口分布均衡度水平,二是细化的人口与经济分布数据来源(主要是非普查年份的来源)。对问题一,现有研究还没有明确提出这样的均衡度指标(有不均衡程度的指标,如上文提到的份额偏离程度等)。对问题二,已有研究主要来自地理学界,如田永中等(2004)基于土地利用信息进行的人口密度模拟;卓莉等(2005)基于夜间灯光遥感进行的全国人口密度分布模拟;王露等(2014)基于分县尺度数据对2020~2030年的全国人口分布进行的预测等。这些研究大多依赖于庞大的基础数据,但很多数据不便获取。

对于上述问题,本文尝试提出一个测度人口分布均衡水平的指标,并设计一套以数量不太大的节点和常规统计数据为基础的 Matlab 空间分布模拟方法,从而能够比较简便地实现对人口与经济分布的模拟及对人口分布均衡水平指标的估算。

## 二、模拟方法及理论基础

人口或经济分布的空间模拟需要有一定数量的空间节点作为基础来进行插值计算,如果纯粹从模拟精度来看,无疑节点越多、空间分布越密集,得到的结果精度越高。但是,中国幅员辽阔,高密度的空间节点采集从人力、物力、耗时等各方面来讲都未必适合较高频率(如按年度)的日常统计工作,这只适宜在若干年一次的大规模普查中进行。对日常研究和常规政策制定(精度要求不是特别高的情况)而言,在节省精力和成本的前提下迅速得出正确结论则显得更为重要。省级数据显然太粗,但地级、县级行政单元数量又很多(超过 300 个地级行政单元、超过 2 000 个县级行政单元),仍会带来较大的工作量,并且受各种因素影响总有一部分地级和县级行政单元的统计数据难以完全收集齐。在考虑以上相关因素后,一个可行的折中办法是选择部分节点(包括较为重要的关键节点和部分容易获取数据的其他节点)作为基础,在一定的理论指导下,用合适的方法对非节点区域进行插值计算,从而得到一个可以正确反映真实情况(但非完全精确)的结果。从便于获取数据及降低工作量等角度出发,选择的节点数量不宜过大、涉及的指标不宜太多、太复杂,插值计算的方法也应该比较简单和易于操作。

### (一) 插值计算方式的选择

根据本文的研究目的,我们拟选取部分地级以上行政单元作为节点,将全国 31 个省份范围按地理经纬度坐标进行人口与经济密度的网格化插值,然后通过全部网格化的数据来比较每个插值点上的人口与经济密度情况。当我们模拟出研究期内各年(2005~2012 年)的人口与经济分布情况之后,通过二者的比较可以了解全国人口分布是否逐渐变得更加均衡。

各种空间分析软件中提供了多样化的插值方式,以本文使用的 Matlab 软件为例,其中较为常用的插值方式包括最邻近插值(Nearest)、线性插值(Linear)、三次插值(Cubic)和 Matlab 提供的独有插值方法(v4)等。不同的插值方法隐含了基于节点值的不同算法,在节点值一定的前提下也可能会带来差别较大的插值结果,因此对插值方式的选择非常重要。通过对比常用的各种插值方式,我们认为线性插值是相对更容易控制的一种方式,它直接按照两个节点之间的指标值落差进行等距离变化插值,不会因为软件自身的算法问题而产生不可控制的插值结果。但只有本身就是线性变化的变量才能直接使用线性插值方式,而对于非线性变化的变量,则需要进行转换。

现有研究几乎均表明人口或经济密度的距离衰减并不是线性的,因此需要构建出一个关于人口或经济密度的新替代变量,该替代变量随距离增减呈线性变化。于是可以把原始密度指标值先换算成这个新替代变量,再代入 Matlab 软件中进行线性插值运算,而运算出来的结果则需要反过来重新换算成原始的密度指标变量才能进行人口与经济实际分布的分析。

(二) 基于相关理论的线性关系构建

现有相关研究中采用的距离衰减函数有多种,比较常用的为 Clark 模型、Gauss 模型等。不过这些模型只是实证经验模型,没有得到充分的理论证明(特别是缺少经济学角度的理论和运行机理分析)。陈彦光(2000)对相关模型的理论推导做了有益的探索,但对其中一些参数取值的经济理论依据没有进行充分探讨,不能很好地指导本文进行空间模拟。尹虹潘(2005)及其后续研究基于引力模型提出经济中心的吸引强度,即:

$$E=kQ^{\alpha}/d^b \tag{1}$$

其中, $E$  为经济中心在某区位的吸引强度, $Q$  为经济中心的“质量”(可用人口规模或生产总值等来体现), $d$  为该区位与经济中心间的距离, $k$ 、 $\alpha$  为大于 0 的常数, $b$  为距离衰减指数。 $b$  的表达式为:

$$b=2/3\mu \tag{2}$$

其中,参数  $\mu$  为城镇化率(城镇常住人口占总人口比重)。随着城镇化率提高,指数  $b$  将趋于降低,也即经济中心的吸引强度衰减会趋缓。以此为基础进行类似分析,以经济总量为研究对象,也可以知道参数  $\mu$  就应该指经济非农化比例(第二、三产业的增加值占地区生产总值的比重)。

人口或经济密度的距离衰减本身就是经济中心吸引强度随距离衰减的重要体现,因此人口或经济密度值  $\rho$  应该是关于经济中心吸引强度的增函数,这里假定其具体函数形式为:

$$f(\rho)=\lambda\rho^{\beta}=E \tag{3}$$

其中, $\lambda$  和  $\beta$  都是大于 0 的参数,至于具体取值下文再讨论。

根据式(1)至式(3)可以得到两个待插值点  $i$  和  $j$  的人口或经济密度的关系:

$$\frac{\lambda\rho_i^{\beta}}{\lambda\rho_j^{\beta}}=\frac{kQ^{\alpha}/d_i^{2/3\mu}}{kQ^{\alpha}/d_j^{2/3\mu}}=\frac{d_i^{-2/3\mu}}{d_j^{-2/3\mu}} \tag{4}$$

通过系列简单变换之后可以得到一个新的替代变量  $\rho^{-3\mu\beta/2}$ (密度  $\rho$  的函数),这个新的替代变量满足按距离呈线性变化的规律,可以作为替代原始密度变量代入 Matlab 进行线性插值运算的新变量,即:

$$\frac{\rho_i^{-3\mu\beta/2}}{\rho_j^{-3\mu\beta/2}}=\frac{d_i}{d_j} \tag{5}$$

至此,我们得到了密度距离衰减函数中参数  $b$  的取值依据,也构建出了满足随距离增减而呈线性变化的新替代变量(关于密度的函数),为空间模拟找到了基本理论支撑。

由于参数  $\lambda$  在构建的线性关系当中已经消掉,因此不是我们关心的重点,那么不妨假设  $\lambda$  为常数,再来探讨  $\beta$  的取值。

由式(1)和式(3),并根据经济中心的“质量”(人口规模或生产总值) $Q$  等于经济中心的人口或经济密度  $\rho_0$  与经济中心面积  $s$ (没有行政区划调整时可以把  $s$  作为常数处理)的乘积,容易得到以下关系:



$$\beta \ln(\rho) = \ln(k/\lambda) + \alpha \ln(\rho_0 s) - b \ln(d) \quad (6)$$

令式(6)中的全部常数项之和等于  $c$ , 则式(6)可转化为:

$$\ln(\rho) = \frac{1}{\beta} [c + \alpha \ln(\rho_0) - b \ln(d)] \quad (7)$$

随着经济的不断发展,  $b$  将不断减小, 而  $c$ 、 $\alpha$  均是常数,  $\rho$ 、 $\rho_0$  是否同向变化, 以及如果同向变化时谁又会变化得更快则是不确定的, 这要取决于经济中心是处于偏向以集聚化为主要特征的极化状态还是偏向以扩散化为主要特征的辐射状态。倪鹏飞等(2014)的研究表明, 人口或经济重心很可能是围绕中心城市(如省会)呈聚集和扩张交错进行的脉动性状, 因此要一般性地判断  $\rho$ 、 $\rho_0$  的变化关系是难以做到的。根据其他参数的取值情况, 即使在  $d$  不变的前提下, 我们仍不能判断  $\beta$  是否为常数, 以及如果它不是常数时的变化趋势。但根据经验判断,  $\beta$  的取值很可能在 1 附近。鉴于此, 在模拟过程中, 我们首先将按照  $\beta=1$  来进行初步计算, 再根据模拟计算的结果与实际统计值之间的差距, 将  $\beta$  值作为对初步模拟结果进行微调修正的工具参数。

### 三、数据来源及模拟设定

#### (一) 关键节点选取

本文选择的关键节点分为两类: 第一类是对全国人口与经济分布具有重要影响的 36 个城市, 其中包括直辖市(4 个)、省会(自治区首府)城市(27 个)、非省会的副省级城市(5 个)。第二类关键节点主要用于加大关键节点密度、提高模拟精度, 在重点考虑空间分布等因素的前提下, 从全国的地级行政单元中选取了 53 个地级市(地区、州、盟), 保证除北京、上海、天津 3 个直辖市及宁夏回族自治区之外的其他省份中至少有 2 个及以上的关键节点, 对于辖区面积大、内部地区差异大的省份适当增加关键节点数量, 关键节点数量最多的省份中达到 7 个节点, 但省际交界地带原则上不同时在相邻省份重复选择关键节点。两类关键节点的总量为 89 个, 已经达到一定数量, 并且空间分布相对比较密集, 能够满足模拟分析的需要(见表 1、图 1)。

#### (二) 控制节点选取

除了关键节点之外, 我们还设定了相关控制节点作为辅助。控制节点分两类: 第一类是插值区域的四角, 按照中国版图所处经纬度, 我们设定东经 70~140 度、赤道至北纬 60 度之间的区域作为插值区域, 由于 Matlab 软件的线性插值是一种内插值方式, 所以四角作为控制节点是必要的, 但四角距离相对较远, 其对插值结果的影响可以忽略不计。第二类是基本可以看做人口与经济密度为 0 的点, 根据经验, 主要是从辽宁丹东出发沿逆时针方向到广西西南部的内陆国界线, 这一段国界线沿线区域地广人稀, 近似设定为“无人区”参与辅助插值应该是大致符合现实的。此外, 在川藏交界处、藏疆青交界处基本也可以近似认为是“无人区”, 两处各选择 1 个控制点。对所有的控制点, 人口密度和经济密度原则上都可以设

表 1 模拟插值中选取的关键节点

省份	关键节点城市	省份	关键节点城市
北京	北京	湖北	武汉、襄阳
天津	天津	湖南	长沙、张家界、郴州
河北	石家庄、秦皇岛	广东	广州、深圳
山西	太原、朔州	广西	南宁、桂林、北海、贵港
内蒙古	呼和浩特、呼伦贝尔、赤峰、锡林郭勒、巴彦淖尔	海南	海口、三亚
辽宁	沈阳、大连、丹东	重庆	重庆、万州、黔江
吉林	长春、四平、松原	四川	成都、广元、宜宾、达州、雅安、阿坝、凉山
黑龙江	哈尔滨、齐齐哈尔、大庆、佳木斯	贵州	贵阳、遵义
上海	上海	云南	昆明、曲靖、丽江、红河、大理
江苏	南京、连云港、盐城	西藏	拉萨、昌都、日喀则
浙江	杭州、宁波、温州	陕西	西安、安康
安徽	合肥、黄山	甘肃	兰州、嘉峪关、金昌、庆阳
福建	福州、厦门	青海	西宁、玉树、海西
江西	南昌、赣州	宁夏	银川
山东	济南、青岛、东营、烟台	新疆	乌鲁木齐、吐鲁番、哈密、巴音郭楞、和田
河南	郑州、安阳、信阳		

注：重庆直辖市面积相当于一个中等省，在选取的 3 个关键节点中，重庆节点按 1997 年前原四川省重庆市所辖区域口径处理，以便与其他关键节点匹配，万州、黔江为市辖区。



图 1 关键节点空间分布示意图

为 0,但考虑到将使用负指数形式的密度替代变量参与计算,所以实际模拟时的密度值设定为  $1^{-10}$ ,可以认为其近似等于 0。

(三) 数据来源及处理方式

研究目标决定了我们希望使用的人口基础数据是常住人口数据,而非户籍人口数据。研究期选择从 2005 年开始,是因为 2005 年后全国的常住人口统计体系更加健全,即使不能直接获取地级行政单元的常住人口,但基于常住人口口径的其他相关指标(如人均地区生产总值等)也符合研究的需要。

按照前面提出的指标尽量简单的原则,用于插值计算的指标主要是各节点的人口密度(每平方公里人口数)和经济密度(每平方公里增加值),用各年度《中国区域统计年鉴》中 89 个关键节点的常住人口和地区生产总值数据除以其辖区面积得到人口密度与经济密度值。其中 2010 年前的人口指标为户籍人口口径,用来进行人口分布模拟不合适,而人均

地区生产总值是按常住人口口径计算的,所以我们采取用地区生产总值除以人均地区生产总值的方法得到常住人口的近似值。虽然这样处理并不完全准确,但在缺乏直接常住人口数据来源的情况下不失为一种可以接受的替代处理方式。2010年之后发布的人口数据本身是常住人口口径,因此可以直接使用。《中国区域经济统计年鉴》里少数年份中有个别关键节点统计指标缺失,我们主要通过关键节点所在地区的统计资料来补齐。

2005~2012年各年的全国城镇化率指标、经济非农化(第二、三产业增加值比重)指标等可以通过各年的《中国统计年鉴》或国家统计局的“国家数据”网站<sup>①</sup>获得。但需要说明的是,由于地方的人口数据加总并不等于全国人口总数,地方的地区生产总值加总也不等于全国的国内生产总值(GDP),而我们使用的关键节点数据都是地方数据,所以这里的全国城镇化率指标、经济非农化指标使用的是中国31个省份相应指标直接加总计算的数值,以便与89个关键节点的数据口径更加匹配。

所有89个关键节点的经纬度数据,我们通过国家测绘地理信息局的“天地图”网站<sup>②</sup>对相应节点进行手工取点获取。控制节点(主要是边界线)的经纬度坐标通过Matlab软件读入包含中国国界线信息的shp文件提取。

#### (四) 模拟分析中的其他设定

插值区域及网格化精度设定。研究的对象区域设定为31个省份(不含港、澳、台地区),因此将东经70度至140度、赤道至北纬60度之间的整个区域,按每1度经度及每1度纬度都等分为10个网格点的插值精度进行插值,网格化为700(东西方向)×600(南北方向)的数据点矩阵。

各类节点指标赋值及范围控制处理方式设定。关键节点采用上文方式处理得到的人口密度和经济密度分别进行人口分布与经济分布的插值模拟计算;控制节点也按上文的方法赋值参与插值计算。对于广西到辽宁丹东的海岸线沿线城市,先在不设控制条件的情况下进行插值,但这样必然会出现插值溢出到海岸线之外的情况,我们将通过Matlab的边界识别算法删除掉出现在海岸线之外的数据点。

插值后回算验证的相关处理方式设定。对插值结果,我们拟通过回算全国人口总量和经济总量并与实际统计值进行对比的方式进行验证。在插值时是按经纬度每1度等分为10个插值点来处理的,插值完成后回算总人口和GDP总量时,将假定全国区域内的每个插值点代表着一个以其为中心的小方格区域,用人口或经济密度与小方格区域的面积相乘即得到该小方格区域的人口数量或生产总值。但地球表面是一个球面而非平面,因此实际上存在北半球两条既定经线之间的纬线段(东西向)距离随纬度越高而不断缩短的问题。也就是说,不同纬度上的小方格南北向的距离相等,而东西向的距离随纬度增加而减小。我们按照

① 国家统计局网站网址: <http://data.stats.gov.cn>。

② “天地图”网站网址: <http://www.tianditu.cn>。

球面距离算法计算了东经 70 度、东经 140 度两条经线上从赤道至北纬 60 度之间每隔 5 度的纬线段长度(东西向共 13 个距离值),并求出这 13 个距离值每两个相邻者的平均值(共 12 个),然后可以非常近似地把插值区域视为宽(南北向的边)相等、长(东西向的边)由南向北不断缩短的 12 个矩形共同构成的平面区域,而 12 个矩形区域包含的插值点(及其代

表 2 插值区域不同纬度东西距离及面积当量

西界点坐标		东界点坐标		纬线段长度	平均长度	面积
东经(度)	北纬(度)	东经(度)	北纬(度)	(公里)	(公里)	当量
70	60	140	60	3710.45	3993.39	0.51
70	55	140	55	4276.34	4546.53	0.58
70	50	140	50	4816.73	5071.96	0.65
70	45	140	45	5327.19	5565.10	0.72
70	40	140	40	5803.01	6021.16	0.77
70	35	140	35	6239.31	6435.20	0.83
70	30	140	30	6631.08	6802.22	0.87
70	25	140	25	6973.35	7117.32	0.92
70	20	140	20	7261.29	7375.89	0.95
70	15	140	15	7490.49	7573.83	0.97
70	10	140	10	7657.17	7707.79	0.99
70	5	140	5	7758.41	7775.39	1.00
70	0	140	0	7792.37		

注:“纬线段长度”根据球面距离算法计算,“长度平均”为每隔 5 个纬度的两条纬线段长度的平均值。

表的小方格)数量是完全相同的。以此为基础,如果以赤道与北纬 5 度之间每个数据点所代表的 1 个小方格面积当量作为 1 的话,可以计算出每 5 度之间的单个小方格的面积当量值(见表 2),并可以计算出 1 当量的面积大约相当于 125.9999 平方公里。

四、模拟结果及均衡度分析

(一) 人口与经济分布模拟结果

将选取的关键节点及控制节点的经纬度、人口密度、经济密度数据,按照前面设定的方法用 Matlab 软件的 griddata 函数分别进行人口与经济分布的线性插值运算(假定 $\beta=1$ ),并根据模拟结果回算出全国总人口和 GDP 数据,与实际统计值进行对比。需要说明的是,由于口径和误差等因素影响,各地区数据之和与全国统计值并不完全相等,而关键节点数据都是地区数据,所以这里全国统计值(包括人口和 GDP)使用的是 31 个省份相应指标的直接加总值。

从表 3、表 4 中可以看出,模拟效果较好,通过 89 个关键节点数据在未对人口和生产总值进行其他限定的情况下,完全依据理论推导结论便可以得到与实际统计值总体吻合度非常高的结果。这表明节点的选择及模拟的方法设计较为符合中国的实际。

这个初步模拟结果是在假定  $\beta=1$  时得到的,因此也印证了我们对  $\beta$  值可能在 1 附近的经验判断,并且无论是人口还是 GDP,模拟值与统计值的比率总体上似乎都呈递增的趋



势(逐年的结果并非严格满足这一点,但总体趋势大致如此)。特别是,2009年的人口与经济初步模拟结果与统计值基本一致。

虽然初步模拟值与统计值之间吻合度已经较高,但二者之间仍有正负几个百分点的差距,且人口模拟值与统计值的比率相对经济(GDP)模拟值而言略高。为了更加准确地分析人口与经济分布情况,我们认为有必要引入 $\beta$ 值对初步模拟结果进行微调 and 修正,使模拟值与实际统计值的差距缩小到更充分的程度(可以认为二者完全相等),这里设定的标准是模拟值与统计值的比率在保留两位小数时为100.00%。

由于没有 $\beta$ 值的确切表达式,只能通过逐渐试验来得到使模拟值与统计值近似相等的 $\beta$ 值。初步模拟结果表明, $\beta$ 确实处于1附近,使试验有大致范围可循。结果显示, $\beta$ 值总体呈递增趋势(见表5)。当然,由于 $\beta$ 值是试验反推出来的,其中也可能包含了由于没有引入其他某些因素(如地理、气候等)造成的误差,但这些因素在研究期内基本不会有太大变化,因此可以认为引入这些因素不会从根本上改变 $\beta$ 值的变化方向和趋势。

根据式(3)可知,当控制住经济中心在某区位的吸引强度 $E$ 不变(那么经济中心的密度 $\rho_0$ 也就不变)且 $\lambda$ 被设定为常数时, $\beta$ 值递增意味着该区位的密度 $\rho$ 值是递减的,即 $\rho$ 相对于 $\rho_0$ 在减小。也就是说,经济中心相对于受其影响的区域更呈集聚化趋势,这说明中国在研究期内总体上处在进一步集聚发展的趋势中。

表 3 全国人口模拟值与统计值比较

年份	人口统计值 (万人)	人口模拟值 (万人)	模拟值/统计值 (%)
2005	128604.00	127344.95	99.02
2006	129523.00	128015.72	98.84
2007	130393.00	129734.28	99.49
2008	131434.00	131403.36	99.98
2009	132443.00	132472.73	100.02
2010	133385.00	139069.16	104.26
2011	134042.00	138362.16	103.22
2012	134789.00	140032.01	103.89

注:表中人口统计值为31个省份加总数据。

表 4 全国 GDP 模拟值与统计值比较

年份	GDP 统计值 (亿元)	GDP 模拟值 (亿元)	模拟值/统计值 (%)
2005	199228.10	191482.47	96.11
2006	232836.74	223624.48	96.04
2007	279737.86	268393.44	95.94
2008	333313.96	322776.90	96.84
2009	365303.69	365806.97	100.14
2010	437041.99	438235.08	100.27
2011	521441.11	514320.38	98.63
2012	576551.84	579433.48	100.50

注:表中 GDP 统计值为31个省份加总数据。

表 5 使模拟值与统计值近似相等的 $\beta$ 值

年份	人口模拟 $\beta$ 值	经济模拟 $\beta$ 值
2005	0.9035	0.8748
2006	0.8882	0.8722
2007	0.9545	0.8683
2008	0.9975	0.8940
2009	1.0018	1.0005
2010	1.3135	1.0050
2011	1.2270	0.9520
2012	1.2730	1.0122

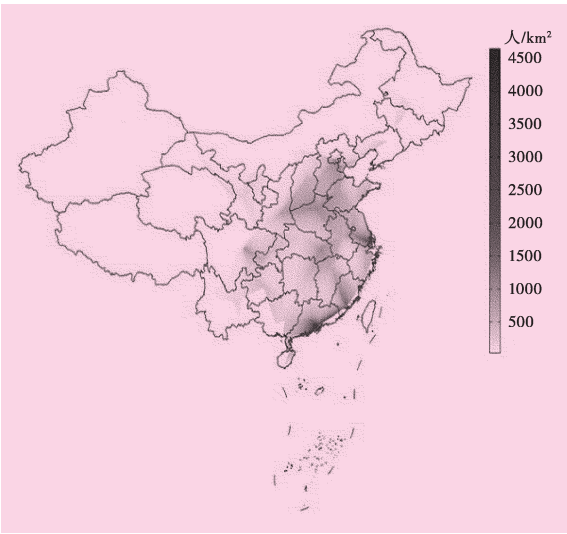


图2 2012年全国人口密度分布

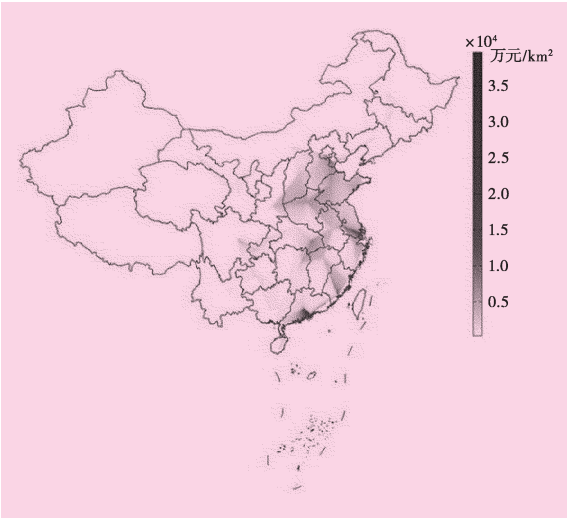


图3 2012年全国经济密度分布

经过加入参数 $\beta$ 微调后的模拟值已经可以视为与实际统计值完全一致。限于篇幅，这里不能一一给出研究期内每年的人口与经济分布图，仅展示2012年的人口密度分布图和经济密度分布图（见图2、图3），从两图可以看出，“胡焕庸线”（瑗瑋—腾冲线）的东南面人口与经济相对密集分布，而其西北面人口与经济密度都较低；经济的集聚程度总体上大于人口集聚度，二者集聚程度的不一致即意味着，相对于经济分布情况而言，人口分布并不是完全均衡的。

### （二）“人口分布均衡度”指标的估算

基于上述对各相关年份全国人口分布、经济分布的模拟结果，我们希望提出一个“人口分布均衡度”指标用于体现人口分布相对于经济分布的均衡水平。从理论上讲，如果将全国每个区位上的生产总值都除以全国的人均国内生产总值（人均GDP），得到人均GDP处处相等时的理想人口分布情况，这便是以经济分布情况为基准的完全均衡人口分布情况。用实际人口分布与完全均衡的人口分布进行比较，可以知道在全国总人口中到底有多大比例的人口已经处于人口完全均衡分布状态下他们应该处所在的位置。这个比例就可以作为判断人口分布总体均衡水平的“人口分

布均衡度”指标。该比例越高表明全国人口分布均衡水平越高，反之亦然。如果全部人口实际所处的位置都与完全均衡分布状态下应该处所在的位置一致，那么该比例即为100%，也就表示实际的全国人口分布本身是完全均衡的。

从模拟的技术实现上，我们以每个插值数据点的生产总值同时除以全国人均GDP的模拟值（经上文微调后可以认为与统计值一致），便得到相对于经济分布而言的人口完全均衡分布情况（即人均地区生产总值处处相同假设下的人口分布）。然后以人口实际分布模拟结果与这个均衡分布情况进行对比，二者的交叠部分占总人口的比重即为人口分布均衡度。

用 100%减去人口分布均衡度估算值，得到的剩余比例的人口就是还没有处于完全均衡分布状态下应该处所在的位置，此时会有一部分区域经济密度相对较高、人口密度相对经济密度而言显得较低（即实际人均地区生产总值较高），另一些区域则恰好相反。

（三）2005～2012 年的人口分布均衡度变化

按照上面提出的“人口分布均衡度”指标及估算方法，得到 2005～2012 年人口分布均衡度如表 6 所示。2005～2012 年，经济分布基准下的人口分布均衡度提高了 5 个多百分点，从整体趋势上看，正在持续变得更加均衡。然而，这个变化趋势是怎么产生的？是人口分布在不断向经济分布靠拢还是

相反？上文指出单纯使用重心分析法有时可能会误导我们，但基于细化分布情况得到结论之后，用重

心分析法从整体趋势上进行旁证却未尝不可，因为细化分布情况已经可以在很大程度上排除了可能形成误导的因素。我们以模拟得到的人口与经济分布情况，计算了 2005～2012 年全国人口重心和经济重心。重心点经、纬度计算方法分别为：

$$Lon_G = \frac{\sum_{i=1}^n Z_i Lon_i}{\sum_{i=1}^n Z_i}, Lat_G = \frac{\sum_{i=1}^n Z_i Lat_i}{\sum_{i=1}^n Z_i} \tag{8}$$

其中， $Lon$  代表经度， $Lat$  代表纬度， $G$  代表重心点， $i$  代表插值点， $n$  为插值点数量， $Z$  为插值点的人口数量或比重（计算人口重心时）、地区生产总值或比重（计算经济重心时）。

计算结果表明，人口重心大致位于湖北省东北部，2005～2010 年总体呈自西向东移动的趋势，但 2010 年以后开始反转向西移动（见图 4、表 7）。而经济重心大致位于河南省东南部与湖北省交界地带，其中 2005～2009 年向西北移动，2009 年后向西南移动（见图 5、表 7）。人口重心和经济重心总体是相向移动的，两重心之间的距离呈缩小的趋势（见表 7），这也

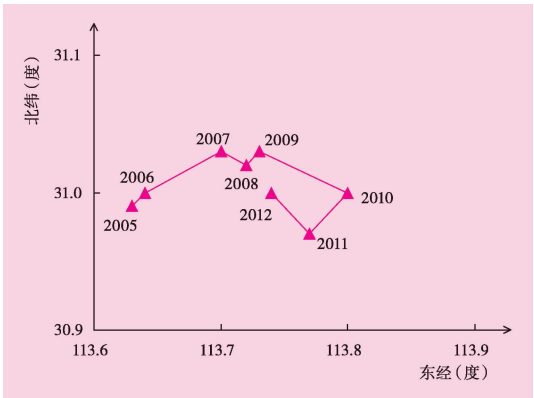


图 4 2005～2012 年全国人口重心移动趋势

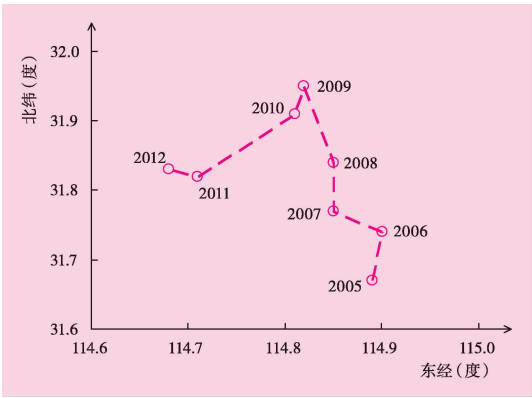


图 5 2005～2012 年全国经济重心移动趋势

表 7 2005~2012 年全国人口—经济重心距离变化

年份	人口重心坐标		经济重心坐标		人口与经济重心 距离(公里)
	东经(度)	北纬(度)	东经(度)	北纬(度)	
2005	113.63	30.99	114.89	31.67	141.86
2006	113.64	31.00	114.90	31.74	145.39
2007	113.70	31.03	114.85	31.77	137.32
2008	113.72	31.02	114.85	31.84	140.58
2009	113.73	31.03	114.82	31.95	145.82
2010	113.80	31.00	114.81	31.91	139.03
2011	113.77	30.97	114.71	31.82	129.69
2012	113.74	31.00	114.68	31.83	128.28

注:人口与经济重心的距离根据球面距离算法计算。

是相对经济分布而言人口分布正在趋于更加均衡的一个表征,可以对上面的人口分布均衡度估算结果形成一定的佐证。

五、结 语

人口和各种经济要素在社会主义市场经济条件下存在着内在的流动与相互匹配的机制,这种机制影响着资源的空间配置、并且在动态协调着人口与经济的空间分布关系。因此,实现相对经济分布而言的人口分布均衡,既是人口长期均衡发展的重要内涵,也是以人为核心的新型城镇化和区域经济协调发展的内在要求。基于上述思想,本文的研究做了以下几方面的探索:(1)提出了以经济分布情况为基准的“人口分布均衡度”概念,将“人均国内(地区)生产总值处处相等时的人口分布”界定为理想的“完全均衡人口分布”状态,而现实人口分布与“完全均衡人口分布”的重叠比例则可以作为判断人口分布均衡水平的指标,奠定了进行定量测度的基础。(2)设计了基于 Matlab 软件的人口与经济分布空间模拟和人口分布均衡度估算方法。与经济学视角的现有相关研究相比,本方法克服了对空间因素考虑不足且大多使用省级数据不够精细等缺陷,使研究的准确性和可靠性有所提高;与地理学视角的现有相关研究比较,本方法的优势在于不需要依赖庞大的地理信息专业基础数据,解决了常规研究中数据不易获得的问题,并且相对更简便的方法可以带来研究时间和成本的节省。当然,本方法的最终效果取决于经济理论指导和空间节点选择,经济理论选择不当、与现实人口与经济分布规律不符可能导致方向性偏差,关键节点选择过少、分布过于稀疏则模拟精确度会打折扣。而这些选择又在一定程度上依赖于研究者的经验和主观判断,不可避免会存在一些人为因素引起的局限。正是为了尽可能减少这些局限,本文设计了通过空间模拟结果回算与实际统计数据进行比对的验证环节,确保研究取得应有的效果。(3)应用本文设计的空间模拟与估算方法,以中国 89 个地级以上行政单元作为节点进行模拟的结果与实际统计值具有较高的吻合度,这说明模拟结果总体上能够正确反映人口与经济现实分布情况,而以此为基础估算出 2005~2012 年以经济分布为基准的“人口分布均衡度”总体上在不断提高,同期的人口与经济重心距离趋于缩短也佐证了这一结论。

基于本文的研究结论,我们提出以下政策建议:(1)区域经济发展理论指出人口与各种经济要素的内在流动机制既包括集聚机制,也包括扩散机制,并且相关研究也表明,集聚和扩散两种机制在长期的区域发展进程中可能是交错进行的,文中  $\beta$  值的变化特征表明研究



期内(2005~2012年)中国人口与经济都正处在集聚发展的趋势中。那么在集聚趋势没改变之前,应该通过不断完善交通等重大基础设施建设来降低物流时间和成本,通过全面深化改革来进一步扫清人口与要素流动的各种制度障碍,按照全国主体功能区规划促进人口与经济向相对适宜地区集聚发展。(2)促进人口与经济的集聚发展。目的是节约集约高效利用国土空间和资源要素,总体方向上应该是引导欠发达地区的富余人口在市场力量的作用下自愿向经济集聚程度高于人口集聚程度的相对发达地区转移,在更高发展水平上实现人口与经济分布均衡。但在集聚发展的总趋势下,也需防止个别一线城市过度集聚带来负面影响,其周边的重点城市及城市群应逐渐分担一线城市的部分人口与经济承载功能。(3)研究期内全国经济重心持续西移及2010年后人口重心西移的趋势,虽不排除有珠三角区域的影响(该区域部分城市处于人口或经济重心以西),但也在一定程度上表明中西部内陆地区的部分重点中心城市应成为未来集聚发展的承载区域。因此,在继续促进沿海地区率先发展的同时,也应结合新时期国家“一带一路”(丝绸之路经济带、21世纪海上丝绸之路)和长江经济带发展等战略实施,适度扶持内陆部分重点中心城市及其周边的城市群,以此带动中西部地区协调发展和沿海—内陆对外开放联动格局的加快形成。(4)政府的公共资源配置要更加立足于以人为核心理念,顺应人口流动的内在规律和未来人口分布格局持续演变需要,打破过去主要以行政层级高低、行政区域单元为依据的固化配置方式,转而在建设用地指标供应、基础设施建设、社会保障体系建设和公共服务供给等方面,形成围绕常住人口数量和结构变化这个中心环节而与之适应的公共资源动态配置新机制,为促进人口分布均衡提供有力保障。

#### 参考文献:

1. 陈彦光(2000):《城市人口空间分布函数的理论基础与修正形式——利用最大熵方法推导关于城市人口密度衰减的 Clark 模型》,《华中师范大学学报(自然科学版)》,第4期。
2. 封志明、刘晓娜(2013):《中国人口分布与经济发展空间一致性研究》,《人口与经济》,第2期。
3. 郭斌、文雯(2013):《中国人口与经济分布地域差异及格局演变》,《经济地理》,第2期。
4. 胡安俊、刘元春(2013):《中国区域经济重心漂移与均衡化走势》,《经济理论与经济管理》,第12期。
5. 倪鹏飞等(2014):《经济重心与人口重心的时空演变——来自省会城市的证据》,《中国人口科学》,第1期。
6. 田永中等(2004):《基于土地利用的中国人口密度模拟》,《地理学报》,第2期。
7. 王露等(2014):《基于分县尺度的2020~2030年中国未来人口分布》,《地理研究》,第2期。
8. 尹虹潘(2005):《对城市吸引区范围界定的理论分析》,《财经研究》,第11期。
9. 张车伟、蔡翼飞(2013):《人口与经济分布匹配视角下的中国区域均衡发展》,《人口研究》,第6期。
10. 卓莉等(2005):《基于夜间灯光数据的中国人口密度模拟》,《地理学报》,第2期。

(责任编辑:朱 犁)