

我国实际抚养比研究

——2023-2037年的预测

张 苏 朱梓怡

【摘 要】用理论抚养比来分析我国真实的劳动力负担时会导致很大的偏差。为解决该问题,在实际抚养比计算公式中引入了更合理的年龄段,加入了劳动参与率以及不同年龄段的消费需求,还创新性地引入系统聚类法将全国划分为4个区域,严格按照普查年鉴中每4岁一个年龄段统计人口的特点,对GM(1,1)模型各期预测值的重要性参数进行了合理调整,最终得到了2023-2037年4个区域的人口及少儿、老年抚养比预测值。

【关键词】理论抚养比;实际抚养比;系统聚类;抚养比预测

【作者简介】张苏(1975-),男,教授,中央财经大学经济学院(北京 100081);朱梓怡(通讯作者),北京师范大学统计学院,zzycufe@foxmail.com(北京 100875)。

【原文出处】《天津大学学报》:社会科学版,2023.5.421~425

【基金项目】国家社会科学基金项目(21BJL037)。

在人口红利逐渐消失,老龄化逐渐显现的背景下,国家出台了一系列鼓励生育的政策。自2016年1月1日起,国家放开二胎政策,2021年三胎政策落地,但政策效果的显现需要时间。为应对人口老龄化问题,当前需要密切关注的一个指标是抚养比,也就是非劳动人口与可劳动人口的数量之比。这一指标有助于了解劳动力的抚养负担,关乎如何设计可持续的养老比政策。

估算抚养比通常的做法是采用统一国际标准的公式进行计算,并没有考虑劳动参与率、人口消费水平等因素,计算出来的抚养比与劳动力真实负担有巨大差异。

近年来,传统的抚养比测算方法导致的问题引起了学者们的广泛关注。肖周燕^[1]通过第五次人口普查资料发现,理论抚养比和考虑了劳动参与率的实际抚养比之间有非常大的差距。根据李中秋和王朝明^[2]的观点,名义的抚养比(即理论抚养比)更侧重反映的是人口的年龄结构,其测算社会实际劳动力的能力弱。晏森等^[3]提到,理论抚养比的劳动人口与

非劳动人口年龄段规定不符合我国法律法规。毕蔚兰^[4]研究认为,名义抚养比虽然是国际通用的测算抚养比的方法,但这其实是一种简化的方法,实际的抚养比还需要考虑就业状况等因素。

由于这些原因,学者们致力于引入新的指标和方法,以克服理论抚养比在社会真实劳动力负担衡量方面的缺陷。陈涛等^[5]探索了人口抚养比如何转变为社会抚养比,该研究认为理论抚养比在经济角度上能够很好反映社会年龄结构的变化,有助于测算“过剩人口”变动,但为了能更准确地反映社会实际抚养比情况,还需要加入人口素质与劳动参与率等指标^[6-7]。姚引妹^[8]认为理论抚养比夸大了劳动年龄人口规模,这是由于我国法律实际上不允许14岁以下的少年儿童就业,因此理论抚养比所规定的第一个阶段,即0~14岁实际上低估了少儿人口,因此对3个年龄段的划分节点重新进行了规定,以平均初次就业年龄(19岁)和平均退休年龄(58岁)对年龄段进行重新划分。由于劳动参与率这一指标会对理论抚养比与实际抚养比情况的偏离造成显著影响,部

分学者对如何引入劳动参与率进行了有益的探索^[2,9]。比如,通过劳动年龄人口数减去在校学生及关押人口数来近似估算实际的劳动人口;引入有效劳动人口的概念用以替代劳动年龄的人口等。

根据本文的研究,实际抚养比可能比理论抚养比高出十倍有余,本文在进行抚养比预测时(未包含港澳台地区),充分考虑了上述学者提出的问题,对抚养比模型进行了3个修订:对劳动年龄的划分重新进行界定、将劳动参与率纳入考虑并进行地域划分来研究其抚养比差异。

一、考虑劳动参与率与不同年龄段消费需求的抚养比测算模型

理论抚养比指的是每100名劳动人口需要负担多少名非劳动人口(包括老年人口与少儿人口)。我国当前计算抚养比采用国际通用的抚养比计算公式,理论总抚养比(GDR)等于理论少儿抚养比(CDR)与理论老年抚养比(ODR)之和。其中,劳动年龄、少儿与老年的年龄规定分别为15~64岁、0~14岁及65岁以上。

如前文所述,由于未进行更合理的年龄段划分、未考虑劳动参与率因素,致使理论抚养比与社会真实的抚养比差距很大。本文的计算模型将有如下3个方面的改进。

一是在劳动年龄划分方面,需要考虑普查资料上每4岁一个小年龄段的特点,应当就近取整,以确保人口计算的准确性。

二是考虑劳动参与率因素。

三是研究抚养比的地域差异。

由于按照规定划分的东北、西北、华北、华中、华东、华南、西南各地域间的省份只是地理位置上具有相似性,在人口的各方面指标上,不一定具有相似性。因此,本文通过聚类方法划分区域,这是指标相似度最好的划分方法。确定划分原则后,先使用国际规定的方法计算出预测的抚养比(理论值),再重新考虑劳动年龄、劳动参与率等因素后,计算出实际的预测抚养比。

1. 劳动年龄的界定

国际上一般规定,劳动人口年龄为15~64岁,但我国不允许16岁以下人口就业。另外,根据我国规

定,6岁及以上青少年需要完成九年义务教育,初中毕业时,平均年龄应为15岁,还不可以就业。从高质量就业角度看,专科、本科毕业生是就业主体,可参照青少年就业的平均年龄为21.5岁^[10]。

由于人口普查资料中是从0岁开始,以4岁为一阶段对人口进行统计,因此本文将就业年龄约等为20岁。另外,退休平均年龄约等为59岁。

2. 劳动参与率的计算

根据CEIC数据库中的数据,自2009年起到2020年,我国劳动人口参与率从71.743%逐年下降到67.465%。我们研究发现,全国平均劳动参与率随年份变化基本呈现逐年匀速递减的趋势,每年的劳动人口参与率相较于上年递减0.373%,整体上反映了我国的人口红利在逐渐消失、老龄化逐年严重的趋势。本文采用一元线性方程对我国劳动人口参与率进行拟合,计算得到未来20年我国的平均劳动人口参与率 $R(t)=63.8\%$ 。

3. 考虑劳动平均参与率因素后的劳动人口

未来各年各区域的实际劳动人口数为

$$P'_{20 \sim 59}(t) = P_{20 \sim 59}(t) \times R(t) \quad (1)$$

式中, $P'_{20 \sim 59}(t)$ 是第 t 年20岁到59岁实际参与劳动的人口,而 $P_{20 \sim 59}(t)$ 是第 t 年20到59岁的总人口数(设定2015年为基期,记此时 $t=0$, $R(t)$ 是未来20年的平均劳动人口参与率)。

4. 考虑不同年龄段的消费需求

国际标准规定没有考虑不同年龄段人口间消费的差异性。我国实行的义务教育保障青少年在义务教育阶段免除学费,且部分学校还有餐饮食宿补贴,总体上青少年的消费相较于成年人和老年人更少。另外,老年人的消费主要在医疗卫生方面,虽然我国目前医保体系日渐完善,也出台了相关大病以及多项药物的补贴政策。但是总体来看,独居的老年人的总消费支出仍不可忽视。本文提出假设:0~19岁青少年的消费需求为成年人的0.5倍,考虑到双减政策影响,教育培训费大幅减少了,而60岁以上老年人(不包括60岁)的消费设定为成年人的1.5倍。

5. 实际少儿及老年抚养比

由以上分析可以得到,各地区实际的少儿、老年抚养比如下

$$CDR' = \frac{P_{0-19}(t)}{2P'_{20-59}(t)} \times 100\% \quad (2)$$

$$ODR' = \frac{3P_{60+}(t)}{2P'_{20-59}(t)} \times 100\% \quad (3)$$

式中, $P_{0-19}(t)$ 与 $P_{60+}(t)$ 分别表示该区域 0~19 岁和 60 岁及以上的第 t 年人口数。

6. 基于灰色预测 GM(1,1) 模型对分年龄段的人口预测模型

在预测未来人口抚养比时,在多项式法、时间序列和 GM(1,1) 这三种方法中,用 GM(1,1) 是最佳模型,预测精度最高^[11-12],本文采用改进后的 GM(1,1) 模型进行预测。

本文认为,原始 GM(1,1) 模型中赋予 $x^{(1)}(k)$ 和 $x^{(1)}(k-1)$ 同等重要的权重值得商榷,在时间序列指数平滑模型中,距离当期越近的过去数据比重越大,本文认为距离原始数据越近的预测数据比重应该越大。而预测期数越高的数据再去预测其余数据时,比重应该减小(未来不可控因素很多,因此预测期数越高,其预测不准确的可能性越大)。根据指数平滑模型中平滑系数的确定方法,我们将公式修改为

$$x^{(0)}(k) + a[(1-m)^k x^{(1)}(k) + (1-m)^{k-1} x^{(1)}(k-1)] = b \quad (4)$$

其中 m 类似于指数平滑模型的平滑系数,由于人口增长速度并不会剧烈变动。参照指数平滑模型的取值,在本文中,对于 0~19 岁及 20~59 岁这两个年龄段将 m 取值为 0.05, 60 岁及以上年龄段 m 取值为 0.01。这是因为 60 岁及以上老年人口数总体来说需要更多考虑已知数据,否则最终会与实际老年人口数偏离越来越大。最后使用最小二乘法对 $\hat{k} = \begin{bmatrix} a \\ b \end{bmatrix}$ 进行估计,并进行预测。

二、抚养比预测

1. 使用层次聚类法对人口特征进行区域划分

我们选用样本省份分年龄段指标(包括分年龄段死亡率、分年龄段生育率、分年龄段性别比、分年龄段人口数)作为各省份的特征变量,使用系统聚类法对其进行聚类(本文未将港澳台数据纳入计算)。

聚类过程中,我们试验了最短距离法、最长距离法、中间距离法、Mcquitty 相似法、类平均法、重心法与离差平方和法后,认为最好避免某个省份单独成类的分类方法是最长距离法,因此我们使用了最长

距离法先将样本分为七类:区域一为黑龙江;区域二为广东;区域三包含河南与山东;区域四包含四川、河北与江苏;区域五包含贵州、北京、上海、内蒙古、吉林、甘肃、重庆、青海、海南、新疆、西藏、宁夏与天津;区域六包含湖北、浙江、湖南与安徽;区域七包含山西、陕西、福建、辽宁、广西、江西与云南。可以看出,区域一至区域四的每个区域包含的省份很少,说明将样本划分为 7 个区域,区域数偏多,有必要进行合理合并。

接下来对聚类结果进行合理调整。根据文献,一般将聚类结果划分为 4~5 类较为合适^[13-14]。根据数据的聚类结果,将全国 31 个省市自治区及直辖市(未包含港澳台地区)分为 4 类时,可以使省份单独成类的情况最少,且各类别中省份个数较为合适。区域一为黑龙江省;区域二为广东、河南、山东、四川、河北、江苏;区域三为贵州、北京、上海、内蒙古、吉林、甘肃、重庆、青海、海南、新疆、西藏、宁夏、天津;区域四为湖北、浙江、湖南、安徽、山西、陕西、福建、辽宁、广西、江西、云南。下文的抚养比预测模型表明,这一聚类结果是有意义的。

2. 对各区域抚养比的计算及预测

(1) 人口预测。本文统计了各省市自治区 2000 年、2005 年、2010 年与 2015 年人口普查及 1% 人口抽样的调查数据^①,使用聚类划分的各区域内指标的平均水平,对 2023~2037 年共 15 年的人口数用 GM(1,1) 方法进行预测。

对于划分的 4 个区域,其人口预测结果可以概括为 3 种类型。

首先是迅速衰老型。区域一(黑龙江省)的人口模式,可概括为迅速衰老型,其 60 岁及以上老年人口将在 2032 年超过青壮年人口。根据预测,截至 2037 年,区域一 0~19 岁人口预计达到 52 万人,20~25 岁人口预计达到 6928 万人,60 岁及以上人口约为 1 亿人,老年人口数量占全部人口比重约为 60%,区域一将面临劳动力严重不足且劳动力人口负担非常沉重等问题。另外,其少儿人口占比非常低,未来青壮年劳动力极其缺乏,全省将面临快速衰老的严重问题。

其次是较慢衰老型。区域三(贵州、北京、上海、内蒙古、吉林、甘肃、重庆、青海、海南、新疆、西藏、宁

夏、天津)的人口增长模式可以归纳为较慢衰老型,其60岁及以上老年人口的占比正逐年加大,从2023年计算的27.6%增长至2037年的47.1%。截至2037年,老年人口数量达到2446万人,没有超越20~59岁年龄段人口2616万人,因此其人口的衰老情况相较迅速衰老型模式缓慢得多。

最后是老年与青壮年人口增长率平衡型。区域二(即广东、河南、山东、四川、河北、江苏几个省份构成)和区域四(湖北、浙江、湖南、安徽、山西、陕西、福建、辽宁、广西、江西、云南)可以被归于一类。根据预计,区域二各省0~19岁的平均人口将从2023年计算的1442万人逐年下降到2037年的1052万人。同时,60岁及以上人口将从3602万人增长到1.2亿人;区域四的劳动人口发展情况与区域二有共性,该区域20~59岁即青壮年人口增长率与60岁及以上老年人口增长率基本持平,在短时间内不会出现劳动力严重缺失的情况,但0~19岁的平均人口与区域二一样呈现出逐年下降的趋势,由于劳动力补充不足,区域二和区域四的各省都难以长久保持劳动力与老年人口数量增长持平的情况。

(2)抚养比预测。本文分别按照式(2)和式(3)对各区域的实际抚养比进行计算,分别得到各区域的抚养比。

迅速衰老区域的抚养比。区域一(黑龙江省)2037年的实际老年人口抚养比(ODR')较2023年增长了200%还多,老年人口抚养比高达366.1%,而其实际少儿抚养比(CDR')始终是所有区域中最低的,最高时也未超过5%。另外,无论是少儿抚养比还是老年抚养比,其实际与理论值的整体趋势都大致相同,但同年度的理论与实际老年抚养比数值相差较多。

较慢衰老区域的抚养比。区域三(贵州、北京、上海、内蒙古、吉林、甘肃、重庆、青海、海南、新疆、西藏、宁夏、天津)的人口模式被归为较慢衰老型。虽然其预测结果同样是少儿抚养比逐年递减而老年抚养比逐年上升,但相较迅速衰老区域,区域三CDR'在2023年的起始值较高,为12.0%。同时,其2037年ODR'的估计值为219.8%,相较迅速衰老型区域低很多,且15年间的增幅较小,增幅为113.0%。

老年与青壮年人口增长率平衡区域的抚养比。区域二(即广东、河南、山东、四川、河北、江苏)老年人口占比在2023年初始情况为20.8%,比区域一高。同时,其2023年初始青少年人口占比为8.6%,远大于区域一。在2023年至2037年这15年间,其青壮年人口增长率为168.3%,老年人口增长率为264.7%,该区域的老年人口与青壮年人口增长率差距是4个区域中最小的,能够达到暂时性的增长平衡。从其理论值与实际值的比较中可以看出,区域二CDR的2023年估计值相较该年CDR'高出近10%,说明在不考虑劳动参与率等情况时,该区域的少儿人口占比相较其他区域要高。

区域四(湖北、浙江、湖南、安徽、山西、陕西、福建、辽宁、广西、江西、云南)也属于老年与青壮年人口增长率平衡区域。其ODR'在2023年时预计为117.1%,2023年到2037年的15年间,区域四的ODR'仅增长了56.8%,该区域CDR'在2023年的初始值与区域三相近,为10.8%。区域四的CDR起始值为17.02%,说明其少儿人口占比较高,然而随着时间推移,其少儿人口数量下降较快,劳动人口补给渐渐不足,老龄化不断加重。

三、结语

由于通用的理论抚养比没有考虑劳动参与率等因素,以此计算真实的劳动力负担将会有偏差,本文设计了新的模型,引入更合理的年龄段,加入劳动参与率以及不同年龄段的消费需求,并进行了系统聚类法计算,将全国划分为四大区域,整合出了三个抚养比类型。第一,迅速衰老区域。2023-2037年间,该区域ODR'增长了219.3%,CDR'降低了86.2%。第二,较慢衰老区域。2023-2037年间,该区域ODR'增长了113%,CDR'降低了66.7%。第三,老年人与青壮年人口增长率平衡区域。2023-2037年间,两个区域的ODR'分别增加了35.9%与56.8%,CDR'分别降低了72.8%和74.0%。

我国通用的是理论抚养比,按照国家统计局数据^②,2021年我国ODR为20.8%,2007年至2021年这15年间全国的ODR增长率达到87.4%。根据本文对比,理论公式计算出的抚养比与实际抚养比有显著差异,且ODR远低于ODR',因此根据本文估计,实际

老年人口抚养比增长率在未来15年会更高。可见,我国老龄化的程度被低估了,应迅速采取更有效的人口政策以及养老金政策。

注释:

- ①数据来源:各省统计局官网。
②数据来源:国家统计局官网。

参考文献:

- [1]肖周燕.理论老年抚养比与实际老年抚养比偏离分析[J].人口研究,2004(3):38-43.
[2]李中秋,王朝明.我国人口老龄化过程中实际抚养比与抚养比偏离率的估算[J].理论与改革,2015(5):55-59.
[3]晏森,王高社,杨克俊.我国人口抚养比指标探讨[J].陕西行政学院学报,2008(3):36-38.
[4]毕蔚兰.名义老年抚养比与实际养老负担的偏离分析[J].环球市场信息导报,2016(22):26-27.
[5]陈涛,陈功,宋新明,等.从人口抚养比到社会抚养比的探索分析[J].中国人口科学,2008(2):24-32.

[6]铁瑛,张明志,陈榕景.人口结构转型、人口红利演进与出口增长:来自中国城市层面的经验证据[J].经济研究,2019,54(5):164-180.

[7]田巍,姚洋,余森杰,等.人口结构与国际贸易[J].经济研究,2013,48(11):87-99.

[8]姚引妹.人口抚养比:理论与实际的偏离及修正[J].中国人口科学,2010(6):2-12.

[9]郝乐,吴頔.人口抚养比的精确化[J].沈阳大学学报(社会科学版),2017,19(5):552-555.

[10]孙萍.新疆南疆三地州人口抚养比的预测分析研究[J].乌鲁木齐职业大学学报,2018,27(4):10-16.

[11]王倩,熊德斌.人口抚养比变化趋势研究[J].中国集体经济,2018(8):157-158.

[12]吴天魁,王波.上海市社会总抚养比的预测分析:基于最小二乘法的改进GM(1,1)模型[J].应用泛函分析学报,2014,16(4):374-379.

[13]高慧娟,薛玉莲,方忠.基于系统聚类的各省金融发展现状比较[J].福建师范大学学报,2021,39(2):189-195.

[14]徐业钊,张晗,冯贯昂,等.基于系统聚类法的基金资产配置策略研究[J].中国商论,2021,840(17):102-104.

Research on the Actual Dependency Ratio in China: Forecast for 2023-2037

Zhang Su Zhu Ziyi

Abstract: Using theoretical dependency ratios to analyze the real labor burden in China leads to significant deviations. To solve this problem, the calculation formula of the actual dependency ratio was introduced, incorporating a more reasonable age range, labor participation rate and consumption demand of different age groups. The hierarchical clustering method was innovatively introduced to divide the country into four major regions and key parameters of the GM(1,1) model's predicted values for each period were adjusted strictly according to the characteristics of population statistics for every four years in the census yearbook. Finally, the predicted population and child as well as elderly dependency ratios for the four regions from 2023 to 2037 were obtained.

Key words: theoretical dependency ratio; actual dependency ratio; hierarchical clustering; dependency ratio prediction