

为什么 PPP 还需要做重大修正?

——评霍诺汉《使用 PPP 进行国家间比较的优势与不足》

邱 东

(江西财经大学统计学院,江西 南昌 330013)

[摘要] 霍诺汉先生新近关于 ICP 的论文《使用 PPP 进行国家间比较的优势与不足》(Using Purchasing Power Parities to Compare Countries: Strength and Shortcomings)深入探讨了国际比较机理和经济意义,并指出了“购买力平价(PPP)”的不足,其批判精神与思考内容均值得重视。本文是对霍诺汉论文的一个经济统计学评论^①,基于对论文背景及内容的概述,探讨了文中值得关注和思考的几点问题,主要包括:ICP 基于对市场汇率法的否定;各轮 ICP 间的“动态一致性”;基于生产率因素对 ICP 结果的修正;注重 ICP 的比较机理和现实意义;基于全球化因素对 ICP 的修正,并进一步地对未来 ICP 需要拓展研究的若干课题提出展望。经济统计在全球经济事务中起着独立作用,对于国际经济统计标准与数据,我们不仅要认真解读和应用,同时也要深入挖掘其内在的测度、核算和比较逻辑,以弥补和修正现行标准的不足,强化国际经济治理。

[关键词] 购买力平价(PPP);国际比较项目(ICP);市场汇率法;生产率;ICP 修正

[DOI 编码] 10.13962/j.cnki.37-1486/f.2022.04.002

[中图分类号]F019 **[文献标识码]**A **[文章编号]**2095-3410(2022)04-0017-13

一、霍诺汉论文的背景和概要

彼得森国际经济研究院(the Peterson Institute for International Economics,PIIE)2020 年第 16 期《政策概要》(Policy Brief)发表了霍诺汉(Honohan)先生的《使用 PPP 进行国家间比较的优势与不足》(Using Purchasing Power Parities to Compare Countries: Strength and Shortcomings)^[1]。

2009-2015 年,霍诺汉先生曾担任爱尔兰中央银行的总裁和欧洲中央银行治理委员会的成员,他也是彼得森国际经济研究院的外籍院士。仅从其社会身份看,霍诺汉先生似乎并不是经济统计学领域的专家,而是一位资深的经济统计数据用户。

从其阐述内容和观点看,这是近年来国际统计界少有地偏重于经济统计的基本思想,而非

[作者简介]邱东(1957-),男,江苏张家港人,教育部长江学者特聘教授,江西财经大学统计学院讲席教授。主要研究方向:国民经济核算、经济测度、国际比较、中国统计改革与发展、多指标综合评价方法、可持续发展、宏观调节与区域发展等。

^①霍诺汉先生的论文由北京师范大学李昕教授推荐给笔者,本文初稿 2021 年 1 月 15 日完成。曾发给北京师范大学王亚菲教授、李昕教授、王岩副教授,东北财经大学谢长博士、黄雪成博士,国家统计局国际统计中心首席统计师王金萍女士等征求意见,2021 年 8 月 21 日完成第二稿修改。

国际比较方法的技术细节的论文。作者对国际比较机理和经济意义探讨颇深,还在标题中直言不讳地指明“购买力平价(PPP)”的不足。业内罕见此例,其批判精神和思考内容都值得重视。ICP 操作者应该充分重视用户的反馈意见,特别是对 ICP 方法论的“外部冲击”,而不应该满足于圈子内部的方法细节讨论。

霍诺汉先生开篇指出,良好的国际经济政策要求对国民经济运行具备良好可比性的数据,要求采用对经济分析具有“充分意义(meaningful)”的方式。需进一步明确的是,意义容易具备,但达成“充分”却非常困难。“充分意义”的达成绝非易事,至少需要长期的、艰难的专业探索。

霍诺汉先生对 ICP 结果提出了三项重大修正,分别基于生产率因素、全球价值链因素和环境因素,就充分说明了这一点。^①国际比较项目(ICP)已经实施 50 多年了,已成为国际经济统计的一项常规性操作,为什么 PPP 还需要做重大修正?

霍诺汉先生提出,尽管新 PPP 数据是国际经济比较的非常有价值的工具,但在使用时还需要小心,应该理解其局限。霍诺汉先生的论文涉及全球比较的大背景,也数次提及对中国国势不同于 ICP 结果的判断,或者说从另一个角度的深刻解读,这对我们高质量发展的战略与策略构思更具参考价值。霍诺汉先生认为,PPP 调整汇率的功能发挥不及预设,还只是将数据置于跨国可比基础的第一阶段。这个基本判定是全时段的,既基于 ICP 的过去和当下,也着眼于其未来发展。这篇论文提出了三个重大修正。

就 ICP 自身而言,霍诺汉先生特别强调生产率因素对国际比较的影响,需要就此做出重大修正。这篇论文并没有就事论事,而且还专门设置了“方法论事项(methodology issues)”一节,作者特别注重对 ICP 比较机理的剖析,注重对数据现实意义的追究,通篇都做了较为深入的论述。

就 ICP 的经济测度基础而言,因为国内生产总值(GDP)本身的缺陷可能会严重影响到国家间的比较,而 ICP 恰恰以 GDP 作为比较基础,所以我们在使用中应该注意到 ICP 数据质量的基础性缺陷。除生产率因素外,霍诺汉先生还提到了两个重大修正,一个是环境因素,一个是全球化生产链。要保证 ICP 结果的稳健性,就需要考虑这两个重要因素。

就与现有核算体系的关系而言,这两个因素大有不同:前者侧重于“经济福利(economic wellbeing)”,而非“物质福利(physical wellbeing)”,其修正超越了国民核算体系(SNA);而后者则主要涉及国际比较的基础宏观指标,究竟应该是 GDP,还是国民总收入(GNI)? 涉及收入的国际分配问题,但修正思路仍然在 SNA 框架之内。

二、ICP 基于对市场汇率法的否定

ICP 的构建从一开始就基于对“市场汇率法(the MER method)”的否定,尽管多年来“世界银行图表集法(the Atlas method)”作为对市场汇率法的一种改进,通过修匀市场汇率,以解决其剧烈波动带来的比较偏差,取得了一定的效果。但随着 ICP 的日趋成熟,“世界银行图表集法”就成了一种权宜之计,逐步被边缘化。ICP 方法优于市场汇率法,这成了国际统计界和国际经济界压倒性的刻板印象,似乎并无争议。

霍诺汉先生对市场汇率法也持否定态度,论文从此出发分析汇率测度与 PPP 表现之间的

^①这项研究也表明,证实与证伪并不对称。

差异。论文对汇率法的局限主要^①讲到汇率剧烈波动,这并不完全反映相对价格的变化。霍诺汉先生也主张对汇率做出矫正,但是不应该出现严重失误(serious errors)。影响相对价格变动的因素需要多维度思考,这个表述为后面的生产率因素修正埋下了伏笔。

霍诺汉先生考察了汇率波动剧烈的阿根廷等六个国家^②。如果将汇率修匀,即便以 ICP 结果作为标准,PPP 与汇率二者之间也不过是水平之差或程度之差,而变动趋势(方向)大致相同,这意味着 PPP 结果或许可以通过其他方式得到,或许存在另外一种可能性,未必一定需要否定汇率法而另起炉灶。尤其考虑到两大类操作的成本效益对比,ICP 未必在总体上占优。

从计算性质上看,ICP 就是一个平均法,将其结果与汇率观察值放在一起比较波动性大小,并不公平。如果对汇率观察值做平滑处理,再看二者在数据平稳性上的优劣,似乎方法优劣比较的结论才更为令人信服。本来统计方法的长处恰恰是处理剧烈波动的数据,仅从方法论的这个角度看,汇率波动对国际比较未必构成颠覆性的威胁。而且,如果汇率法的结果仅仅存在波动性缺陷,那么由此可以得出的一个逻辑推断就是:如果采用“世界银行图表集法”可以切实消除剧烈波动的影响,那么就可以放弃 ICP。

ICP 的优势需要以巨大投入为代价。霍诺汉先生指出,收集 ICP 价格数据的范围日趋扩展,这是一项巨大的全球和各国家的统计努力。而在 2005 年以前,ICP 的实施往往是“小本经营”,各种统计投入跟不上工作需求,其数据结果的质量也往往不尽如人意。

工作拓展与成本增加是一枚硬币的两面,我们不可能也不应该只取其一。在国际比较布局时,还应该考虑到 ICP 实施的社会成本,当然也要看到 ICP 的社会效益,比如“数据基础结构(data infra-structure)”的构建。还有一点,对 ICP 的功能要求越多,就越容易隐含更多概念和操作上的方法论问题。总之就评价任何一种方法而言,效益和成本都需要等量齐观。

三、各轮 ICP 间的“动态一致性”

各轮 ICP 数据的更新总是伴随着争论,一个典型的例子就是印度和中国这样的发展中大国,其 PPP 数据大幅度地变动:2005 年陡然向下^③,而 2011 年则逆转向上。2017 年第 9 轮基本上维系了 2011 的方法原则^④,其结果也就维持了 2011 年 ICP 的数据指向,但不同轮次的比值差异程度不同。尽管方法论探讨还在接续,现用方法似乎是被大家接受合理的折中方案。

2005 年 ICP 的结果或许出乎人们意料,毕竟这一轮是第一次真正全球意义上的国际比较,而且,1993 年国际比较项目遭受了颠覆性失败,这是时隔多年之后 ICP 重整旗鼓的首次操作。

欧美经济学界好多人不大相信 2005 年 ICP 的某些数据结果,如果将这种结果倒推到 20 世纪 50 年代,印度数百万人的收入水平就会“低于存活边缘水平(levels below subsistence)”,

①霍诺汉先生同时也指出:汇率管制和汇率盯住制度可能造成市场汇率背离相对价格的长期平衡关系。

②其实,还应该考察其他经济体的汇率与 PPP 的关系,才能全面评估二者在数据稳定性上的优劣。然而在 ICP 必然优于 MER 的教条之下,这方面研究往往被忽略。

③正如霍诺汉先生本文所言,2005 年的估计导致几乎所有发展中国家收入和贫困水平的大幅下调,中国和印度的人均收入分别为以前的估计的 40%和 35%。

④迪顿教授和施莱尔先生在其 2020 年关于 ICP 的 NBER 工作论文指出,这是尽可能保持比较方法一致而刻意谋求的结果。笔者提请 ICP 数据用户注意,“形式一致性”与“内容一致性”存在差别,笔者在《深入探索 ICP 隐含的经济测度问题——评〈GDP、福利和健康:2017 年轮 ICP 的若干思考〉》中做了论述。

2005年的ICP数据对穷国价格水平估计过高,而对实际经济规模估计偏低。

对穷国国民而言,这种所谓过低的国际比较结果反倒相对容易理解和接受。坦白而言,富国国民自身的低收入生存能力比较差,即便是富国的穷人,其低收入条件下的生存能力也不如穷国的穷人。虽然同在地球村也恍如隔世,富国多数人也不知道穷国穷人在极低收入条件下如何生存,对这种残酷的社会现实缺乏切身理解。一个典型例证是改革开放初期,中国派到欧美国家的访问学者,国内发的生活津贴比较少,甚至有的低于所在国(发达国家)的贫困线水平,然而,这些访问学者在国外的生活并不比当地穷人差,而且在一年访问到期后回国时,还能给家里节省出“买大件”的钱。认知缺陷往往导致对策失误。由于缺乏对穷国客观环境的真正体会,国际组织和欧美专家还常常对穷国发出类似“何不食肉糜”的政策建议。

就印度和中国等发展中国家的真实规模而言,2011年ICP结果相比于2005年出现了相当大的逆反。这极大可能出于针对2005年结果的人为调整。毕竟对ICP数据的“倾向性管理”还比较容易实现:只要尽量剔除穷国的高价格项目,将该类数据处理为奇异值^①,就可以拉低穷国的价格水平,并放大穷国的实际经济规模,从而避免专家所认定的2005基年比较结果的偏误。

真正全球意义的ICP从第7轮开始,至今只有三轮(即2005年、2011年和2017年,)需要深入思考的是,估计结果与方法论争议之间是否存在某种关联效应?在好多ICP数据用户眼里,2005年似乎“高估”了穷国的价格水平,从而“低估”了穷国的实际经济规模,2011年轮的大反转难道不存在为提升数据质量的人为调整?

可以从方法论逻辑推演得到的是,在新的比较过程中,完全可以采取某些针对性的预防措施,以矫正前期似乎存在的“低估”或“高估”倾向。如果结论先行,事先就可能具有某种数据结果倾向,在方案设计时预设,在过程操作时预调,比较结果究竟可能包含多大的人为因素?值得深究。再者,如果把经过预调的结果当成客观数据,再用其论证所用比较方法的正确性,是否隐含了某种循环论证的味道?

还值得注意的是,2011年的国际比较结果又对贫困统计形成巨大冲击。霍诺汉先生指出:所估计的低收入国家价格水平大大低于原来的认知,相应地,这些国家的平均实际收入水平就特别高,在这些国家,低于世界银行极端贫困线的人数下降了一半。Deaton(2001)^[2]认为,国际比较再次威胁贫困统计,到了“毁坏贫困估计”的地步。世界银行对该项比较结果的反应是,迅速将绝对贫困线由每天1.25元调整到1.90元。鉴于不同基准数据间的大幅波动,阿特金森教授建议,放弃用ICP估计的新结果调整国际贫困线,但此议并没有被采纳。

应该看到,这种数据结果的矛盾客观上提出了一个一般性的问题:究竟如何解决不同经济统计项目之间的冲突^②,或者说,如何保障不同经济统计项目间的协调性?能否将某一种统计操作当作绝对标准?这至今仍然是个有待深入探讨的经济统计方法论问题。

^①剔除所谓奇异值就需要事先界定其定义域,其中不可避免地包含了价值判断。例如,在亚太地区比较中,中国大陆的商品价格如果超过香港便被认定为奇异值,无法进入基础数据录入系统,这种识别程序的依据便是所谓的“宾大效应”

^②这里是ICP与贫困统计之间的矛盾。

四、基于生产率因素对 ICP 结果的修正

在国际标准的基础上比较平均生活水平和贫困发生率是一个标志性的应用,似乎 ICP 不可或缺。然而霍诺汉先生认为,PPP 并非总是适合于此。近来的趋势是将以 PPP 计算的 GDP 作为各经济体相对规模和优势的指标,比如,作为一个因素评估地缘政治力量。例如,就以 PPP 计算的 GDP 而论,2017 年中国刚刚超过美国。但是,霍诺汉先生指出,这种计算忽略了这样一个事实:PPP 包含着较贫困国家价格较低的系统性趋势,如果这种趋势归因于低收入国家在国际贸易项目生产中的低下生产率^①(Balassa, 1964^[3]; Samuelson, 1964^[4]),那么 PPP 的调整就“过头了(goes too far)”,将更多的经济优势归于这些低收入国家,超出了其应当具有的份额。仍以中国为例,如果消除这个系统因素,重新进行经济规模排位,中国将列在美国和欧盟(尽管英国不在其中)之后。如何看待霍诺汉先生提出的生产率因素这个重大修正?

(一)这种 ICP 结果“可调整性”^②说明不同国家间的价格差异是多重原因形成的

霍诺汉先生指出:采用 PPP 方法所剔除的,不仅是汇率误定价造成的扭曲,还可能出于其他因素,其中非常重要的就是各国间的生产率差异。这就是说,穷国的价格水平之低还另有其因,因此不宜将其一概推算为其“实际经济规模(real economic size)”较大。这正是霍诺汉先生提出进行生产率因素调整的机理所在。

相对于高收入国家而言,低收入国家通常平均工资低下,非国际贸易品的价格也低。霍诺汉先生指出:价格水平总体上随着人均 GDP 的增加而提升,其系统作用相当可观,被半个世纪以来的 ICP 数据所证实。Hassan(2016)^[5]的研究表明,这种模式在收入水平低下时不大明显,但当收入上升时证据就愈加充分。Cheung 等(2017)^[6]的研究表明,另外三个因素强烈地影响者价格水平的跨国模式:该国是否为石油输出国?该国在腐败水平是否排在高位?该国是否地理遥远?这些因素每个都倾向于拉升发展中国家的价格水平。

一种对这种模式的理论解释被引用得最多,被概括称为“巴拉萨-萨缪尔森效应(the Balassa-Samuelson effect)”,与供给条件(supply condition)相关。按照这个理论,国家间相对工资水平系统地与生产率水平相关,若某经济体的平均劳动生产率低下,则其工资水平也比较低。国际竞争通常会使得国际贸易品^③的价格趋于相同,但就非国际贸易的货物和服务^④而言,其价格低下反映了低收入国家的低工资率。

需求条件也可能与此相关,霍诺汉先生在论文注释中列示了几项相关研究。De 等

^①Balassa 和 Samuelson 1964 年对此分别做出了经济解释,被广为引用。

^②ICP 数据结果“是否应该调整”与“用什么方法进行调整”是两个层次的问题,前者涉及对国际比较基本概念深层次解读。“来自生产率的价格差异”在经济意义上是实际的,而非名义的。世界银行的 PPP 值也经过了生产率调整,尽管只是针对部分项目和部分国家,也表明了 ICP 数据结果的可调整性。

^③研究“国际贸易项(实物与服务)”对国际价格变动的间接影响同样重要,仅仅关注实物产品还不够,还应该关注服务项。

^④需要明确的是,“国际贸易项”与所谓“非国际贸易项”之间并不存在一堵“柏林墙”。在全球化背景下,二者的相互影响更为密切。但是,“非国际贸易项”受“国际贸易项”影响的内在机制并没有得到深入研究。至于汇率变动是否包含了这种相互影响,也值得探究。当我们将“国际贸易项”与“非国际贸易项”加以区分时,对好多研究者而言,就人为地割断了其间的隐含联系,且不自知。

(1994)^[7]的研究表明,如果消费者对非国际贸易服务的偏好随着收入增加,服务的相对价格将被拉升。Bhagwati (1984)^[8]的研究表明,另一个原因可能是劳动的相对供应,在大多数发展中国家其数量较大,拉低了劳动密集型服务的价格。不过 Devereux (2014)^[9]的研究表明,计量经济研究没有确定哪一种因素占据主导地位。

按照“巴拉萨-萨缪尔森理论(the Balassa-Samuelson theory)”,国际贸易品的平均生产率水平低下,成为低收入国家价格水平低下的系统原因,在这个认知背景下,如果计算以人工汇率(即 PPP)调整的 GDP,剔除该效应,可能会夸大低收入国家的经济规模、生产能力及其在国际关系中的重要性。霍诺汉先生认为,“巴拉萨-萨缪尔森理论(the Balassa-Samuelson theory)”广为接受,具有重要的、却被国际比较忽略的“含义(implication)”,将其纳入 ICP 视野,足以改变各经济体的排位顺序。

从市场汇率到 PPP 应该仅做“部分调整”,换言之,应该计算一个“经过生产率因素调整的 PPP(productivity-adjusted PPP)”,即去掉相对价格变动中由于生产率因素引致的部分。设定如果生产率确实发挥作用的话,则通过将此因素回加到各国的 PPP 估算中,就可以避免各国生产能力测度中的偏差。一个简易概略逼近该份额的方法是估算人均收入对 PPP 的影响,并扣减之。需要强调的是,这里问题的关键在于:所扣减的是反映生产率因素的项目,对数值结果而言,最后究竟意味着加项还是减项,取决于不同国家的生产率水平。

这个调整并不直接标出每个国家的生产率,而只是以人均收入横跨世界增长所表现的平均生产率收益。作为粗略的调整,不如 ICP 那么精细,但却是一种简便方法,揭示了以 PPP 计算的 GDP 的一个重要潜在陷阱,用以比较不同国家经济规模和实力时尤其需要当心。

经过这种调整后,PPP 数据更适于比较不同国家的经济实力,明显地改变了人口大国和地区的排序。霍诺汉先生的论文表现了这种变化,美国仍然处于最高位,中国与欧盟(尽管没有英国)相近,但欧盟排在前面。印度仍然超过日本排在第四,而不是之前的超出 60% 那么多。俄罗斯排在第 6 位,超出“撒哈拉以南非洲(Sub-Saharan Africa)”,不然二者的排位顺序正相反。这些结果表明,没有生产率因素的调整,PPP 也可能误导国家真实经济规模的比较,只不过这种“扭曲”与市场汇率方式不同。

(二)从实用方法角度看,如果可以用生产率因素回调 PPP 数据,这与市场汇率法(或“世界银行图表集法”)的结果相差几何?

从霍诺汉先生的论文可以看出,欧盟和中国按 PPP 计算则都接近美国,而按照经生产率调整的 PPP 计算,GDP 大致占美国的近 80%。但是回顾 2017 年按市场汇率计算的 GDP,欧盟和中国都是美国的 62% 多。可见,考虑生产率因素计算 PPP,从数据结果看是向市场汇率法估算结果的一种回归。

PPP 通常存在着压低穷国价格水平的倾向,也即容易夸大穷国实际经济规模。与此同时,还存在着夸大富国价格水平的倾向,也即容易减缩富国实际经济规模。正如霍诺汉先生指出,PPP 的平均化过程更倾向于体现高收入国家的支出权重,从而 ICP 方法往往会低估低收入国家与高收入国家之间的差距。正是由于 PPP 这种测度和比较失误的陷阱,霍诺汉先生才提出这种重大修正。

需要注意的是,“回调”的是国际比较中误作为价格因素所扣除的部分,是来自生产率因素的影响,而非价格因素自身的影响。国家统计局王金萍女士和张伟先生在《对彼得森研究所 PPP 值调整方法的研究和评价》中指出,“实际上世行发布的 PPP 值也经过了生产率调整,但只是部分区域(亚太地区、非洲地区、西亚地区等)对政府服务支出部分的 PPP 值进行了生产率调整(王金萍、张伟,2021)^[10]。世行调整的理论依据是,政府部分的工资水平不是市场价格,穷国由于占有的资本存量更少,产出效率更低。但在 ICP 实践中发现,一些低工资经济体的产出测算远远超出了正常水平,因此需要调整。”

问题的关键首先在于国际比较中进行生产率因素调整的必要性。北京师范大学王亚菲教授在讨论中提出:“生产率是生产角度的测度工具”,这个视角的揭示非常重要。如果“三方等价原则”基本成立,那么这种调整就是一种经济学意义上的交叉检验,不仅可为,而且非常必要。至于“如何调整”更加符合经济现实,则属于下一层次的问题,有待于深入探讨。

经过生产率调整的 PPP 是一种反向调整,至少会部分抵消压低穷国价格水平倾向的影响。霍诺汉先生提出,PPP 数据使得汇率波动造成的扭曲有效地得到了中性处理。

经过生产率调整的 PPP 是将 PPP 数据往汇率观察值的方向回调,并非那么“中性”。如果说这种调整是“中性”的,那么同时要看到,“世界银行图表集法(Atlas 法)”能够修匀市场汇率波动,也是对汇率观察值的一种“中性”调整。

如果拓宽视野,我们可能面对着两大类四种国际比较数据:PPP 数据、经过生产率调整的 PPP、Atlas 法数据和市场汇率记录值。若将这四种结果列为一个谱系进行比较,那么居中的两种数据才更为“中性”。

不宜先行将“市场汇率法”打入冷宫,或许经过调整,汇率观察值还可以提供有用的信息,最起码可以用作交叉检验,也还是一种专门的信息利用。我们应该进行两大类方法结果比较,看看是否存用其他方法进行国际比较的可能性。尤其是,充分考虑从事 ICP 的成本和效益,我们究竟应该在哪些方面投入经济统计和国际比较的各种资源?

(三) 生产率差异与产出质量差异之间是什么关系?

国际比较是从空间角度计算实际经济规模,基本算式是从以本国货币计算的“名义产出”剔除不同国家间的价格差异,即剔除价格因素所代表的“水分”——名义值与实际值之间的差,从而得到以共同货币单位 PPP 计算的“实际经济规模”。在不同国家间“经济同一产品”的确认中,不可能找到完美的匹配品,只能进行“近似项(the like item)”之间的价格比较^①,事实上将“近似项”处理为“同质项(the same item)”。由于忽略了“潜在质量因素”,很容易将其归结为价格水平因素,或价格影响,才导致生产率差异被人为地剔除。

邱东(2018)^[11]较为系统地阐述了产品质量因素对 PPP 结果的影响,认知方向与霍诺根先生倡导的生产率调整一致,彼此印证。

(四) 霍诺汉先生将欧盟作为一个经济体参与比较,值得国人应该深刻反思

我们从 2012 年 GDP 总量超过日本开始,就自封为“世界上第二大经济体”,诸多国际格局

^①Compare the like with the like,而非 compare the same with the same。

分析中根本没把欧盟当作一个“测度单位”和“比较单位”,显然,持这种见解的中国专家还缺乏大国竞争的博弈智慧和定力。应该看到,自从柏林墙倒塌之后,大国竞争的主战场就是美欧博弈,这是中国快速发展的一个重要外部因素,故而应该是中国进入全球化发展深入“知彼”的一个重要内容。

五、注重 ICP 的比较机理和现实意义

在 ICP 实施过程中也会提及方法论,但往往是从操作意义上解读和处理,似乎 ICP 方法总体上已经成熟,其比较机理已经贯通,其社会经济意义的实现已经不成问题。不过,霍诺汉先生在论文阐述中却比较注重对方法论问题的探讨。

(一) 如何认识 PPP?

ICP 的核心就是计算一个人造的汇率,霍诺汉先生在论文中称之为“综合汇率(synthetic exchange rate)”,如果这个汇率流行于市场,在平价意义上不同国家的价格将相等。换言之,它将使得各国的平均价格均衡。在概念把握上,霍诺汉先生还强调需要注意综合汇率(PPP)与购买力平价理论的区别,按照购买力平价理论,市场汇率倾向收敛于购买力平价。

实施 ICP,用 PPP 取代市场汇率,使得 GDP 实际价值及其相应的收入和生产数据在一个更为稳健的基础上进行比较,成为国际经济分析工具包的基本构成。

这种替代实际上提出了一个经济统计学的基础性问题:我们在分析现实经济问题时,究竟应该以观察值为准,还是以估算值为准?显然,市场汇率是观察值,而 PPP 是我们人为计算出来的综合价格比率,其中充满了种种假设和估计。通常在经济统计中遵循眼见为实的思维方式,原则上以观察值为主^①,但在国际比较中,PPP 成为一个典型的反例。这是否意味着,在国际经济关系的认知上,人(ICP 专家)比市场更聪明?

(二) 如何全面认识相对价格的影响因素?

一般而言,影响“相对价格(relative price)”变动的因素不止于市场汇率波动,还有系统性的生产率水平差异,等等。ICP 在剔除市场汇率波动影响时,将其他因素也都剔除掉了。当国际比较的重心在于实际经济规模时,这个缺陷的影响尤其明显。

(三) 如何合成不同国家间的价格差异?

按照 Deaton 和 Aten(2017)^[12]的看法,个人带有不同的偏好,且面临不同的价格,比较其福利在理论上近乎不可能。对计算跨国价格差异指数的最佳方法,总是存在着争议。一种实用方法是计算所有国家双边相对价格的平均,一种替代方法是针对一套参考价格和数量框架(人造的参照框架或“平均国”)计算相对数。Neary(2004)^[13]和 Oulton(2012)^[14]的研究表明,不同方法间和方法中的方法论选择都会严重影响所估计的 PPP 结果。

(四) 如何看待 ICP 的价格数据基础?

尽管多年来已经有了很大改进,基本价格数据的搜集仍是 ICP 面临的主要挑战,过程仍不完善。有些产品在发达国家比较常见,但在某些发展中国家无法得到,或者仅仅在那些为富人

^①笔者在 2018 年科学出版社出版的《经济测度逻辑挖掘:困难与原则》中对此问题有过论述,参见该书第 49 页。

和外国游客服务的销售点才能得到。这些产品的价格是否应该包含在比较中,或者采取替代方法?已经采取了高度复杂的系统应对这些问题,但争议仍然存在。例如,Ravallion (2018)^[15]的研究表明,在选项采价时可能偏向于国际贸易品,或许这成为过高标示低收入国家价格水平的贡献因素。

这里需要注意的是,有的情形是现实经济结构本身造成的数据缺失,并不是现实中存在该类事项,只要加大统计投入便可获取相应数据。一棵树长满了果子,全面采摘便有收获。但如果有的树枝无果,却硬性规定每一枝都必须显示其果子的相关数据,实际操作显然无法达成这种严苛的设计要求。

(五)如何看待 PPP 在贫困测度中的应用?

Reddy 和 Pogge (2010)^[16]批评 PPP 基于“平均居民户消费篮子”的做法,他们认为,在定义贫困线时,采用反映贫困者消费的商品篮子比较好。Dabalén 等 (2020)^[17]做了一项事关非洲 16 国的研究,发现就其中多数而言,穷人支付的价格低于平均水平。这里采用市场汇率无所助益,PPP 调整对定义贫困线也不能一锤定音,但是其所考虑的因素却非常重要。

Atkinson (2019)^[18]认为,贫困是一个多维概念,对现金收入的购买力测度并不能完全把握之。需要特别注意的是,“收入”与“收益”的差别,并非仅仅在于“实物支付(转移)”项,邱东 (2018)^[19]列示了七项,对深入分析此问题或有启示意义。

六、基于全球化影响因素对 ICP 的修正

霍诺汉先生倡导的这项调整涉及如何选择 ICP 的基础性宏观指标。在全球化发展的大背景下,跨国公司为避税和充分利用各种地方优惠措施,刻意模糊其地理区位的格局,基于资产转移套利,从而扭曲了各国的 GDP 和 GNI 数据,这使得国际比较难以客观进行。

2015 年,由于跨国公司资产所有权的大规模转入,爱尔兰的实际 GDP 仅一年就增长了 25%,尽管这种指标计算符合国际经济统计准则,但却破坏了 GDP 的常规性应用,对描述爱尔兰的生活水平而言,人均 GDP 不再是一个意义充分的测度。反观这个案例,也说明国际经济统计准则还存在漏洞,需要进一步研究改进。

跨国公司的这种资产转移活动对引资国的影响主要有三种:第一,其上报利润仅影响 GDP 但不影响 GNI。第二,大规模资产的折旧同时增大 GDP 和 GNI,指标名称的 G 代表 Gross,正意味着包含了折旧项。第三,跨国公司将总部移至引资国,尽管公司“未分配利润”归属 GNI,但其股东大多数都是非居民身份。

霍诺汉先生指出,可以用来补救的是 ICP 中 GDP 的主要构成项——“实际个人消费(actual individual consumption, AIC)”,这个指标接近基于消费的经济生活水平测度,而且不受跨国公司转移活动的扭曲影响,ICP 将 AIC 视为一种对“平均物质福利(average material wellbeing)”的测度。如果用 AIC 替代 GDP 进行排序,爱尔兰从全球第 4 降为第 21,从占美国水准的 130% 降为 60%,霍诺汉先生认为这种排序比较符合实际。

另一种应对方法是计算 GNI(*) 指标,剔除由于跨国公司资产转移行动造成扭曲的项目,例如该类资产的折旧、知识产权进口和知识产权交易等。通过这种百分比缩减的项目调整,可以得到一种粗略的国际比较,爱尔兰在欧盟的排位从第二降为第八,与采用人均 AIC 的

结果相近。

霍诺汉先生在论文中提出,绝大多数国家的 GDP 与 GNI 几近相等。但从经济统计学的角度看,“1%的绝对值”这个概念非常重要,从时间比较上看,同一经济体两个总量指标的差异可能还不大要紧,而对空间比较而言,对不同类型国家的“国外净要素收入”或正或负,差异相当可观。所以要避免将 GDP 与 GNI 混同的不良倾向。

这是一个长期被忽略的重要问题,尤其是全球生产链乃至价值链将各种经济体网罗在一起之后,更值得深入探讨。主要发达国家历史上 GNI 通常大于 GDP,优势地位决定了他们对“国外净要素收入”不大敏感,但新兴国家恐怕需要更多地关注这个指标^①,即关注国民总收入 GNI 与国内生产总值 GDP 的差异。

论及基础指标,本来就应该采用 GNI 搞国际比较,ICP 采用 GDP 不过是对经济统计现实条件的一种妥协,并非天然就应该以 GDP 为基础指标。既然是不同货币购买力的比较,当然应该用收入指标而非生产指标作为基础。按照迪顿教授 2020 年在美国经济学会年会上的话说,“我们不能消费非我所属 (we cannot consume what doesn't belong to us)” (Deaton, 2020)^[20]。^② 由于全球化对国际经济关系深度且多轮次的影响,爱尔兰的国势研判不仅不能依赖 GDP,还不得不调整估算 GNI(*) 指标,这个典型案例告诫我们,甚至连 GNI 对国家间收入分配的真实状况都有遮蔽,更何况 GDP,我们决不应该对此掉以轻心,“唯 GDP 论”对高质量国势研判而言是相当深的测度陷阱。

七、ICP 需要拓展研究的若干课题

霍诺汉先生特别强调基于经济福利因素的修正。ICP 基于 GDP,由此,GDP 在测度福利上的缺陷就自然地传导到国际比较当中。霍诺根先生认为,由于高度异质性的存在,其对跨国空间比较的负面影响大大高于对单一国家的时间比较。该论文给出的例子是环境退化。需要特别注意的是,霍诺汉先生强调的这两种修正在可行性上还有很大差别。

问题在于,当下 GDP 和 ICP 需要解决的方法论困境已经不少了,再考虑环境因素,至少在目前还看不到比较令人信服的解决之道。爱尔兰采用两种替代方式,以解决跨国公司资产转移对经济比较的影响,在测度逻辑上达到了进行修正的要求。

在修正方法和思路提出时,更需要注重其可行性。其一,修正方法是否破坏了现行测度的基本逻辑,是否能与现行核算制度保持内在一致性?比如,采用“人力资本”概念,SNA 中“最终消费”的概念还能否成立?是否会沦为一种“中间消耗”?其二,当指标修正成为常规统计项目后,各国是否能负担得起所需要的成本?特别是穷国而言,新的测度与他们的宏观管理是否高度相关?穷国是否有足够的资源和人力持续参与全球宏观数据这种公共产品的生产?此次疫情造成规划中的第 10 轮 ICP 推迟,这个事实应该有助于富国专家更容易理解穷国提供公

^①笔者与北京师范大学王亚菲教授、北京工商大学王春云副教授 2018 年撰写了论文《GDP 不是新兴国家测度和比较国力的合宜指标——基于“国外净要素收入”的国家间分布》,对这个问题做了较为深入的阐述。

^②笔者认为,所谓“三方等价原则”是概括经济总量在生产、分配和使用三个方面在总量上的一致性关系,但不能做绝对理解,将三者完全混同。

共产品面临的窘境——参与国际统计项目受制于资源和数据基础结构。

霍诺汉先生在论文结论部分指出,为世界经济体构建 PPP 涉及许多概念性和实践性问题,不应该低估问题存在的规模(scale)。本文认为,除了基于福利因素的修正之外,当下还应该注意以下问题。

一是只做消费项目的国际比较,效果是否更好? 深入思考“实际最终消费(actual individual consumption, AIC)”的替代性,能否用 AIC 代替 GDP 的国际比较? 或者说,ICP 是否一定要搞 GDP 整体的比较? 还是只搞或重点搞 AIC 的国际比较? 从数据质量角度看,似乎只搞 AIC 更好。GDP 其他构成项目的数据质量和可比性显然不如 AIC,其数据与 AIC 数据混在一起,会使得整个 GDP 的比较质量下降。当然还需要考察,发展中国家的 AIC 数据基础究竟如何?

二是如何看待 ICP 结果的一致性? 是否比较轮次的时间间隔缩短就可以保证 ICP 结果的一致性? 不能误以为只要 ICP 的频率加大,就可以解决所有的方法论问题。霍诺根先生提出的重大修正是否还会出现? 机理挖掘需要我们避免沉醉于方法和技术的细节改进,避免单纯追求程序化的设计和操作,深度的测度陷阱或许还在,需要我们深入、拓展且提升方法论思考。

应该认识到,刻意保持 ICP 轮次间数据结果的动态一致性,其实也是一种“数据管理”。本来国际经济关系的客观变化要求比较方法与时俱进,但为了数据结果在不同轮次 ICP 间保持一致,即可以放弃比较方法的改进,恐怕是一种本末倒置的做法。不宜将“形式一致性”当作“内容一致性”去维护,笔者在《深入探索 ICP 隐含的经济测度问题——评〈GDP、福利和健康:2017 年轮 ICP 的若干思考〉》中就此问题做过论述(邱东,2021)^[21]。

三是 ICP 的参与、数据应用与中国国势的高质量研判。对中国经济统计而言,积极参与 ICP 是坚持开放国策的题中应有之义。而且,如果在新历史阶段推动高质量开放,更需要以国际经济统计项目为数据基础。为此需要关注 ICP 和 SNA 等国际标准的发展动向,中国如何切实跟进? 也需要尽可能准备统计条件。“数据基础结构”不单单是技术,不单单是方法,更是“社会基础结构”的组成,是“软实力”的体现,往往无法通过突击短期内弥补客观存在着的社会差距。例如,如果也需要注意更多地关注“实际最终消费”指标,我们的数据基础如何?

优化对 ICP 数据的应用,前提是对该数据含义的“适当解读(appropriate interpretation)”,同时,对国际比较机理也能有基本的理解。坦白地说,我们在这方面还需要系统地补课。在《量化世界——联合国的主意和统计》中,沃德教授特别强调避免数据的“蒙昧解读(uneducated interpretation)”,我们身处发展中国家,尤其不能掉以轻心。

高质量的国势研判对高质量发展至关重要,对国际经济统计标准和数据,我们既要认真解读和应用,同时也需要学习霍诺根先生的做法,秉持一种科学认知和批判的态度,不能盲目地照搬照抄,而应该深入挖掘其内在的测度、核算和比较逻辑。

总之,ICP 的方法论和数据理解、应用问题总是存在且变化着,需要持续地研究其解决之道,不可能一劳永逸。

参考文献:

- [1] Honohan P. Using purchasing power parities to compare countries: Strengths and shortcomings[J]. Policy

Briefs, 2020, (12): PB20-16.

[2] Deaton A. Counting the world's poor: Problems and possible solutions[J]. The World Bank Research Observer, 2001, 16(2): 125-147.

[3] Balassa B. The purchasing-power parity doctrine: A reappraisal[J]. Journal of Political Economy, 1964, 72(6): 584-596.

[4] Samuelson P A. Theoretical notes on trade problems[J]. The Review of Economics and Statistics, 1964, 46: 145-154.

[5] Hassan F. The price of development: The Penn-Balassa-Samuelson effect revisited[J]. Journal of International Economics, 2016, 102: 291-309.

[6] Cheung Y W, Chinn M, Nong X. Estimating currency misalignment using the Penn effect: It is not as simple as it looks[J]. International Finance, 2017, 20(3): 222-242.

[7] De Gregorio J, Giovannini A, Wolf H C. International evidence on tradables and nontradables inflation[J]. European Economic Review, 1994, 38(6): 1225-1244.

[8] Bhagwati J N. Why are services cheaper in the poor countries? [J]. The Economic Journal, 1984, 94(374): 279-286.

[9] Devereux M B. Real exchange rates and the Balassa-Samuelson effect revisited[J]. NBER Reporter, 2014.

[10] 王金萍, 张伟. 对彼得森研究所 PPP 值调整方法的研究与评价[J]. 调研世界, 2021, (07): 77-80.

[11] 邱东. 国际比较项目基本类别 PPP 中隐含的“纯价比假设”及其经济意义[J]. 经济统计学(季刊), 2018, (02): 38-67.

[12] Deaton A, Aten B. Trying to understand the PPPs in ICP 2011: Why are the results so different? [J]. American Economic Journal: Macroeconomics, 2017, 9(1): 243-64.

[13] Neary J. True multilateral indices for international comparisons of real income[J]. American Economic Review, 2004, 94(5): 1411-1428.

[14] Oulton N. How to measure living standards and productivity[J]. Review of Income and Wealth, 2012, 58(3): 424-456.

[15] Ravallion M. An exploration of the changes in the international comparison program's global economic landscape[J]. World Development, 2018, 105: 201-216.

[16] Reddy S G, Pogge T. How not to count the poor [EB/OL]. https://www.researchgate.net/publication/289994930_How_Not_to_Count_the_Poor/link/56ec0acf08aee4707a384deb/download.

[17] Dabalén A, Gaddis I, Nguyen N T V. CPI bias and its implications for poverty reduction in Africa[J]. The Journal of Economic Inequality, 2020, 18(1): 13-44.

[18] Atkinson A B. Measuring poverty around the world[M]. Princeton: Princeton University Press, 2019.

[19] 邱东. 经济测度逻辑挖掘: 困难与原则(当代经济统计学批判系列)[M]. 北京: 科学出版社, 2018.

[20] Deaton A. GDP and beyond[A]. Summaries from the 2020 Annual Meeting of the American Economic Association[C]. 2020.

[21] 邱东. 深入探索 ICP 隐含的经济测度问题——评《GDP、福利和健康: 2017 年轮 ICP 的若干思考》[J]. 统计研究, 2021, (09): 143-156.

(责任编辑: 路春城)

Why Does PPP Need a Major Revision?

—— Comment on《Using Purchasing Power Parities to Compare Countries:
Strength and Shortcomings》by Honohan

QIU Dong

(School of Statistics, Jiangxi University of Finance and Economics, Nanchang 330013, China)

Abstract: Mr. Honohan's recent paper on ICP 《Using Purchasing Power Parities to Compare Countries: Strength and Shortcomings》deeply discussed the mechanism and economic significance of international comparison and pointed out the deficiency of "purchasing power parity (PPP)", whose critical spirit and thinking content are worth paying attention to. This paper is an economic statistical review of Honohan's paper. Based on the overview of the background and content of the paper, several issues worth paying attention to and thinking about are discussed, including: the denial of ICP based on market exchange rate method; the "dynamic consistency" between ICP rounds; revision of ICP results based on productivity factors; focus on the comparative mechanism and practical significance of ICP; the ICP is revised based on the factors of globalization, and some topics that need to be further studied in ICP in the future are put forward. Economic statistics plays an independent role in global economic affairs. As for international economic statistics standards and data, we should not only carefully interpret and apply them, but also deeply explore their inherent measurement, accounting and comparative logic, so as to make up for and correct the shortcomings of current standards and strengthen international economic governance.

Key words: purchasing power parity (PPP); international comparison projects (ICP); market exchange rate method; productivity; ICP revision

