

# 中国省际人力资本估算 1978 - 2010

——对现有方法的改进

罗 植 赵安平\*

**内容提要** 估算省际人力资本存量对于利用面板数据模型分析中国宏观经济有着重要意义。本文的研究发现，朱平芳和徐大丰（2007）提出的中国城市人力资本估算方法在估计省际人力资本存量时与实际情况存在较大偏差，其直接原因可能是物质资本弹性与人力资本弹性的估计存在应用层面上的缺陷。为此，本文在修改生产函数基本形式与放宽基本假设的基础上，对物质资本弹性与人力资本弹性的设置做出了相应调整，并利用面板数据模型对修正得到的中国省际人力资本存量估算结果进行了检验。研究发现，该数据基本符合实际，可以应用于更广泛的实证研究。

**关键词** 人力资本存量 平均工资 面板数据

## 一 引言

20 世纪 60 年代，舒尔茨（T. W. Schultz）在美国经济学年会上首次提出了人力资本的概念，他认为人力资本是指人们花费在教育、健康、培训、移民和信息获得等方面的投资所形成的资本。随后，Lucas（1988）、Romer（1990）和 Mankiw et al.（1992）等人在经济增长模型中也引入了人力资本的概念，并形成了内生经济增长理论，阐释了人力资本在经济增长中的特殊贡献，由此解释了国家之间的贫富差距以及若干经济现象。因此，人力资本对经济的作用得到了许多经济学家的关注。但是，人

\* 罗植，清华大学社会科学学院和北京市社会科学院管理研究所，电子邮箱：establown@gmail.com；赵安平，中国银行业监督管理委员会北京监管局，电子邮箱：zhaojumping@163.com。

力资本存量没有直接的统计资料可以获取且难以量化,因此准确地对这一指标进行估算十分必要。此外,随着面板数据模型的发展,准确估算省际人力资本存量对利用面板数据模型分析中国宏观经济也有着重要意义。

对于人力资本存量的估算,现有文献从不同角度提出了多种方法。根据出发点的不同,可以将这些方法分为三类:一是投入法,该方法主要从教育资金投入的角度进行考察,如张帆(2000);二是教育年限法,该方法源于Barro & Lee(1993),是目前使用较为广泛的一种方法,廖楚晖(2006)、高慧和余利霞(2007)以及周兴和陈东娇(2007)所做的有关人力资本的研究,都采用了此类方法;三是收入法,这种方法主要从劳动报酬的角度考虑,最早由Mulligan & Sala-i-Martin(1995)提出,该方法本质上是一种将教育与劳动力收入相结合的人力资本测量方法,朱平芳和徐大丰(2007)的估算就是基于这一思路。

虽然现有的估算方法比较丰富,但这些方法在估计中国省际人力资本时均存在一定的缺陷。本文试图通过对现有方法进行修正,得到更加准确的中国省际人力资本存量数据。本文的结构安排如下:第二部分对现有的人力资本存量估算方法进行综述并分析其中存在的不足,第三部分介绍修正后的估算方法,第四部分给出使用修正方法得到的主要估算结果,第五部分是计量经济学检验,最后为本文的结论。

## 二 人力资本存量估算的主要方法及其缺陷

### (一) 投入法

人力资本存量最直接的反映就是形成人力资本所投入的成本,而反映人力资本投入成本的最简单方式,就是国家财政每年在教育上的投入,并由此所形成的教育投入存量。因此,人力资本存量可以通过对教育投入的流量数据使用永续盘存法进行估算(张帆,2000)。这种方法的优点是操作相对简单,但估算结果可能存在很大偏差。其中一个重要原因是教育投入不仅包括国家的财政支出,还可能包括个人的直接投入或间接投入。虽然国家财政支出数据容易获取,但是个人在教育上的投入却难以衡量。另外,使用永续盘存法时,教育投入存量的基值和折旧率也都难以确定。因此,使用该方法一般很难得到较为准确的人力资本存量估算数据。

### (二) 教育年限法

教育年限法考虑了全部人口的受教育程度,主要分为两种。第一种方法是利用流量数据间接估算存量数据。这种方法首先要计算人力资本的流量数据,即每年人力资

本的新增量  $h_t$ ，为各教育阶段的毕业学生中没有继续接受教育的人数乘以他们完成的学习年限的总和，可用如下公式表示：

$$h_t = \sum_i (g_{it} - r_{i+1,t}) y_{it}$$

其中， $i$  表示不同的教育阶段， $g_{it}$  表示第  $i$  个教育阶段的毕业生人数， $r_{i+1,t}$  表示第  $i+1$  个教育阶段的招生人数， $y_{it}$  表示第  $i$  个教育阶段的教育年限。计算出每一年的人力资本增量后就可以利用永续盘存法核算人力资本存量，这其中还涉及选择合适的基年数据和确定人力资本折旧率等问题。

这一方法简明扼要，数据的可得性和准确性都令人满意，许多有关人力资本存量的研究也都使用了这种方法（蔡昉、都阳，2003；廖楚晖，2006）。但是，该方法更适合对全国整体进行估算，而在估算地区人力资本存量时至少存在五个问题。第一，分地区各教育阶段的毕业生人数和招生人数较难获得。第二，人才的流动性难以测量。高等教育是人力资本存量的重要组成部分，但接受高等教育的人往往具有较大的流动性。因此，这种方法可能会高估高等教育资源较丰富地区的人力资本存量。第三，不同时代的教育回报可能不同。第四，各阶段的教育回报也可能不同。第五，“干中学”的因素也会影响这种核算方法的准确性<sup>①</sup>。

第二种方法是直接计算存量数据。这种方法需要首先获得各教育阶段的人数，然后利用下面的公式计算：

$$H_t = \sum_i E_{it} y_{it}$$

其中， $H_t$  为人力资本存量， $E_{it}$  为第  $i$  个教育阶段的受教育人数，而  $y_{it}$  表示该教育阶段的教育年限或回报率。

这一核算方法可以较好解决上一种方法中存在的许多问题，如直接计算存量不受人才流动的影响，使用教育回报率也可以解决不同时代与不同阶段教育回报率不同的问题。但这一方法的主要问题在于可获得的数据比较有限，其中的关键数据全社会人口的受教育情况，虽然可在历年《中国人口统计年鉴》中获得部分数据，但这一数据多来自抽样调查。

### （三）收入法

按照微观经济学理论，劳动力收入反映了要素报酬，因此可以通过测算劳动报酬

<sup>①</sup> 具有相同受教育程度的人，可能会因为工作年限的差别而具有不同的人力资本。

进而衡量人力资本存量,其前提是收入越高的人,人力资本存量也越高。该方法使用劳动者的平均劳动所得来体现劳动者身上所蕴藏的人力资本。虽然这类方法在逻辑上清晰简单,但在实际应用上也可能存在一些潜在问题。第一,劳动者身上所蕴藏的人力资本同他在生产中实际的人力资本供给往往是不一致的,而劳动者在生产中实际供给的人力资本和他应得的劳动报酬也往往存在较大差异。因此,用劳动报酬体现实际的人力资本供给进而计算人力资本存量可能是不精确的。第二,从产出增加值的角度度量劳动者报酬通常是一个汇总的结果,这一结果并未对不同劳动者及其报酬的形式进行分类,而我国从工资、奖金、红利等方面统计的劳动报酬往往低估了劳动者的实际收入,也就低估了人力资本存量的实际价值,使得这种方法的应用效果大打折扣。

比较投入法、教育年限法和收入法这三类估算方法,教育年限法虽相对准确,但可获得的数据十分有限,三类方法中只有收入法有相对丰富且完整的统计资料。Mulligan & Sala-i-Martin (1995) 从单位人力资本的角度建立了人力资本与收入之间的关系,并提出了一种人力资本的估算方法。虽然各种制度因素的存在可能使得不同制度环境下单位人力资本的劳动报酬并不一致,从而使得统计上的收入不能反映真实收入水平,但相同或相似的制度环境下,仍可预期统计上的收入水平可以反映人力资本之间的相对差异。朱平芳和徐大丰(2007)在此基础之上提出了中国城市人力资本存量的估算方法。然而,使用这种方法估算中国省际人力资本存量时,大量的估算数据与实际情况存在明显出入。例如,从人均人力资本存量角度看,利用朱平芳和徐大丰(2007)的方法计算1992年北京市的人均人力资本为11.43<sup>①</sup>,而2000年的人均人力资本相对1992年不仅没有增加,反而下降到7.56。当然,人均人力资本下降可能是由于人口数量增加导致的。但从人力资本存量看,不论按在岗职工数计算,还是按从业人员数计算,人力资本存量的总额也都表现出下降趋势,其中以在岗职工数计算的下降更为明显。例如,使用该方法计算得到的1992年北京市人力资本存量为5446.39,而2000年仅为3009.40,下降幅度超过40%。其原因可能来自两个方面:一是省际数据与城市数据之间可能存在一定的统计差异;二是估算方法可能存在一定问题,导致该方法不适合从省际的角度进行估算。因此,本文主要针对朱平芳和徐大丰(2007)的中国城市人力资本存量估算方法进行调整,进而获得中国省际人力资本存量的估算方法。下一部分将对修改后的估算方法进行详细说明。

① 依据本研究核算方法得到的人均人力资本是一个无量纲数据,只有同方法下得到的数据才具有可比性。

### 三 修正的人力资本存量估算方法

#### (一) 基本框架

首先假设一个规模报酬不变的生产函数。与朱平芳和徐大丰(2007)的假设不同,该生产函数中不仅包括物质资本存量和人力资本存量,还包含一个不可观测的综合变量。这一综合变量包含其他一切可能影响产出的制度因素,其具体形式如下:

$$Y = K^\alpha (AH)^\beta Z^{1-\alpha-\beta} \quad (1)$$

其中,  $Y$  是产出,  $K$  是物质资本存量,  $H$  是人力资本存量,  $A$  代表劳动有效性,  $Z$  代表综合制度变量。他们对产出的弹性分别为  $\alpha$ 、 $\beta$  和  $1 - \alpha - \beta$ 。

之所以引入一个综合制度变量,是因为从实际生产来看,对产出产生影响的因素不可能仅包括资本和劳动力两个部分,制度同样是影响产出与经济增长的重要因素,只是因其通常难以定量测量,因此很难将其放入生产函数中。此外,近30年来中国巨大的制度变迁也被广泛认为是影响人力资本积累的重要因素。在这里,我们以综合变量的形式将其纳入生产函数中。在引入综合制度变量后,我们依然假设生产函数整体是规模报酬不变的,但仅考察物质资本存量与人力资本存量时,则存在非规模报酬不变的可能。

由于人力资本存量可以写成劳动力数量  $L$  与人均人力资本存量的乘积,若假设人均人力资本存量为  $h$ , 则有  $H = hL$ 。因此,可将生产函数表示为人均人力资本的形式:

$$Y = K^\alpha (Ah)^\beta L^\beta Z^{1-\alpha-\beta} \quad (2)$$

假设劳动力的边际收益就是其所获得的报酬,那么,可将劳动报酬表示如下:

$$w(h) = \beta K^\alpha (Ah)^\beta L^{\beta-1} Z^{1-\alpha-\beta} = \beta (K^\alpha / L^{1-\beta}) (Ah)^\beta Z^{1-\alpha-\beta} \quad (3)$$

上式中,劳动者的平均工资  $w$ 、物质资本存量  $K$ , 以及劳动力数量  $L$  都有相应的统计数据,只有劳动有效性  $A$  和综合制度变量  $Z$  没有实际数据。为了处理这两个变量,可假设劳动有效性与综合制度变量同人均人力资本存量之间存在着正相关关系。为了处理上的简化,假设这种正相关关系是线性关系,且当人均人力资本存量等于1 ( $h = 1$ ) 时,劳动有效性为  $A_1$ , 综合制度变量为  $Z_1$ ; 当人均人力资本存量等于  $h$  时,劳动有效性为  $A_h$ , 综合变量制度为  $Z_h$ 。二者之间存在如下关系:

$$A_h = hA_1, Z_h = hZ_1 \quad (4)$$

将此关系代入式 (3) 中, 可分别得到单位人均人力资本存量的工资, 以及人均人力资本存量的工资, 具体如式 (5) 和式 (6) 所示:

$$w(1) = \beta(K^\alpha/L^{1-\beta})A_1^\beta Z_1^{1-\alpha-\beta} \quad (5)$$

$$w(h) = \beta(K^\alpha/L^{1-\beta})(h^2 A_1)^\beta (h Z_1)^{1-\alpha-\beta} \quad (6)$$

用式 (6) 除以式 (5), 可以得到如下关系:

$$\frac{w(h)}{w(1)} = h^{1-\alpha+\beta} \quad (7)$$

由于公式 (5) 中包含了不可测量的变量  $A_1$  和  $Z_1$ , 因此无法计算得到单位人力资本存量。但是, 可以定义一个单位人力资本的效率工资水平  $w(1)^*$ , 令其和  $w(1)$  之间存在如下关系:

$$w(1)^* = \frac{w(1)}{A_1^\beta Z_1^{1-\alpha-\beta}} = \beta \frac{K^\alpha}{L^{1-\beta}} \quad (8)$$

再将式 (8) 代入式 (7), 可以得到人均人力资本存量的计算公式如下:

$$h = A_1^{-\frac{\beta}{1-\alpha+\beta}} Z_1^{\frac{\alpha+\beta-1}{1-\alpha+\beta}} \left[ \frac{w(h)}{w(1)^*} \right]^{\frac{1}{1-\alpha+\beta}}$$

由于  $A_1$  和  $Z_1$  都被假设为常数, 因此, 它们对人均人力资本存量的影响可以被看作是一种单位的变动, 在此将其简化。至此, 可以得到人均人力资本存量的估算公式如下:

$$h_e = \left[ \frac{w(h)}{w(1)^*} \right]^{\frac{1}{1-\alpha+\beta}} \quad (9)$$

在此估算公式中, 劳动力的平均工资和效率工资都可以得到, 只有资本对产出的弹性参数  $\alpha$  和人力资本存量对产出的弹性参数  $\beta$  是未知的。这两个参数对估算结果存在显著影响, 因此需要对其进行必要的说明。

对于参数  $\beta$  的估计, 朱平芳和徐大丰 (2007) 利用式 (2) 和式 (3) 直接得到  $\beta$  的计算公式如下:

$$\beta = \frac{w(h)L}{Y} \quad (10)$$

通过公式 (10) 直接计算参数  $\beta$  在理论上是没有问题的, 但实际计算结果却存在着与理论预期和现实情况差距较大的问题。为了具体说明, 我们使用本研究收集到的

省际数据，并利用公式（10）计算参数 $\beta$ 。其中， $Y$ 使用名义国内生产总值（GDP）数据，而 $w(h)$ 和 $L$ 分别使用在岗职工名义平均工资和在岗职工数。由于计算中存在同单位名义数据相除，并且参数 $\beta$ 也是无量纲的，因此可用名义数据直接计算。通过计算，得到的结果如表 1 所示。

表 1 利用朱平芳和徐大丰（2007）的方法计算部分省份（市、区）主要年份的参数

省(市、区)	1978	1998	2005	2008	2009	2010
北 京	0.1750	0.2195	0.2226	0.2730	0.2945	0.2985
上 海	0.1039	0.1245	0.0993	0.1261	0.1494	0.1417
广 东	0.1888	0.1047	0.0949	0.0915	0.0975	0.0983
山 西	0.1925	0.1487	0.1318	0.1361	0.1472	0.1417
河 南	0.1521	0.1004	0.0920	0.0933	0.1015	0.0971
湖 北	0.1770	0.1236	0.1000	0.0889	0.0998	0.1017
广 西	0.1704	0.1004	0.1019	0.0974	0.1061	0.1015
云 南	0.1902	0.1261	0.1095	0.1209	0.1323	0.1304
新 疆	0.4173	0.1822	0.1421	0.1405	0.1611	0.1501

资料来源：根据作者整理数据计算得到。

表 1 所列的 9 个省份（市、区）分别属于东部、中部和西部。根据通常的划分方法，北京、上海和广东属于东部，山西、河南和湖北属于中部，广西、云南和新疆属于西部。考察上述结果可以发现以下三个特点：第一，从横向上看，东部省份（市）、中部省份和西部省份（区）的结果总体差距不大，比如 2010 年，东部的上海为 0.1417，这一结果还小于西部新疆的 0.1501；第二，从纵向上看，除北京以外，其他各省份（市、区）都有明显的下降趋势，或者是先下降再缓慢增加，比如广东省，从 1978 年的 0.1888 先下降到 2008 年的 0.0915，再增加到 2010 年的 0.0983；第三，西部省份（区）中新疆的估计结果明显高于大部分东部与中部省份（市、区），尤其是 1978 年，新疆的人力资本对产出的弹性高达 0.4173，几乎是上海的四倍。显然，这三点结论与理论预期和中国的实际情况存在显著差异。根据参数 $\beta$ 的经济学意义可知，经济越发达的地方，人力资本存量对产出的贡献率通常应该越高，即参数 $\beta$ 的数值应该越大。因此，朱平芳和徐大丰（2007）在计算参数 $\beta$ 上的方法可能并不适用于省际数据。出现这种结果的原因可能与研究对象自身特点的差异有关，包括但不限于不同层面数据的定义和统计口径不一致的问题，如省内城市间人口流动与省份间人口流动的差异、城市内平均工资与省内城市间平均工资的差异等。

从上面的结果可以发现,通过公式(10)无法得到适用于省际数据的参数 $\beta$ 。为了估计参数 $\beta$ ,我们也尝试了多种方法。最后,我们还是在已有研究的基础上选择假设的方式设置这两个参数。Romer(1996)在其《高级宏观经济学》中指出,发达国家中物质资本的弹性大约是 $1/3$ ,而人力资本存量的弹性会略大于物质资本的弹性,大约介于 $1/3$ 到 $4/9$ 之间。中国是一个发展中国家,从理论上讲,资本对产出的弹性相对于发达国家会偏大,而人力资本存量对产出的弹性相对于发达国家会偏小。同时,章上峰和许冰(2009)在对劳动力弹性的时变估计中,认为中国整体的劳动力对产出弹性从1979年到2005年间维持在 $0.386$ 左右,而资本对产出的弹性大约维持在 $0.614$ 左右。基于以上两个原因,我们从全国整体平均水平的角度对两个参数做出如下假设,假设 $\alpha = 0.60$ , $\beta = 0.35$ 。二者相加并不等于 $1$ ,这是因为生产函数中还包括了难以测量的综合制度变量。这样,在规模报酬不变的假设下,资本和劳动力对产出的贡献都有可能转移到制度因素上,从而降低二者对产出的弹性。至此,可利用公式(8)与公式(9),并基于对参数 $\alpha$ 和 $\beta$ 的假设,估算中国省际人力资本存量。

## (二) 数据选择与来源

数据的选择及其来源同样会对估算的结果产生重要影响。根据上面的估算框架可知,省际人力资本存量的估算主要涉及到三个数据的选择问题,分别是:全社会固定资本存量、劳动力数量以及劳动力所对应的平均工资。下面对这三个数据的选择进行简单说明。

全社会固定资本存量只能估算,并没有统计数据可以获取。许多学者在估算物质资本存量上做出了大量的贡献。由于朱平芳和徐大丰(2007)的研究对象是城市人力资本,没有现成的估算结果可以使用,因此他们在参考了许多研究的基础上,使用了自己的估算结果。而本研究的对象是省际层面,现成的研究成果比较丰富,本研究选择张军等(2004)对省际数据的估算结果,因为该研究不仅综合考虑了众多的研究成果,而且这一结果也得到了许多学者的认同。不过,张军等(2004)并没有给出2006年及以后的估算结果,因此本文使用张军等(2004)的方法进行了估算和补充。其中,由于数据来源问题,无法得到张军等(2004)估算时使用的价格指数。为了处理上的简单,直接使用GDP平减指数来替代该价格指数,而其他的数据来源和计算方式均与其一致。

劳动力数量及其所对应的平均工资有两组数据可以选择。其一是全社会就业人员及全社会平均劳动报酬,其二是在岗职工数及在岗职工平均工资。从模型(2)的意义考虑,劳动力 $L$ 是指全社会就业人员。但是,全社会就业人员的平均劳动报酬这一数



据难以获取，因此只能放弃这组数据，选择在岗职工数和在岗职工平均工资。

除此之外，还有总产出与用于换算实际在岗职工平均工资的价格指数需要确定。总产出使用实际 GDP 衡量。计算方法涉及名义 GDP 与 GDP 实际增长指数。这两个数据主要来自《新中国 55 年统计资料汇编》和历年《中国统计年鉴》。计算方法是先使用迭代法计算实际 GDP 指数，再使用该指数计算实际 GDP。为了处理上的简便，换算实际在岗职工平均工资的价格指数直接使用 GDP 平减指数。

#### 四 省际人力资本存量估算结果

通过上述设定的估算框架和数据，我们对中国省际人力资本存量进行了估算。表 2 中列出了使用公式 (8) 计算的部分地区部分年份的效率工资<sup>①</sup>。

为了便于对比，表 2 列出了三种计算结果：第一，不加括号的结果是用本研究的方法与数据估算得到的；第二，圆括号内的结果是使用朱平芳和徐大丰（2007）的方法，但用本研究的省际数据得到的；第三，方括号中是朱平芳和徐大丰（2007）的估算结果。从表 2 可以看到，三种方法估算得到的效率工资都存在明显的递增趋势，只是不同结果之间的增长速度有所不同。

表 2 效率工资计算结果对比

城 市	1992	1996	2000	2004	2008	2009	2010
北 京	821. 20 (685. 09) [1264]	1244. 80 (1031. 80) [2279]	1860. 40 (2279. 88) [4517]	2353. 19 (2852. 29) [4816]	2364. 42 (1866. 55)	2090. 21 (1357. 73)	2059. 95 (1292. 15)
天 津	635. 37 (492. 54) [658]	842. 76 (835. 17) [1393]	1406. 78 (2119. 12) [2322]	2036. 40 (3889. 93) [3185]	2843. 90 (6463. 04)	2960. 51 (6898. 63)	3296. 99 (8191. 64)
上 海	746. 38 (728. 44) [959]	1201. 92 (1623. 86) [2387]	2061. 08 (4156. 94) [3689]	2865. 19 (7125. 01) [6062]	3654. 46 (9328. 11)	3413. 47 (7563. 17)	3584. 53 (8023. 00)

资料来源：根据作者整理数据计算得到。

① 为了使结果不至于太小，将其扩大 10 倍。这只是一个单位的变化，不会影响具体的估计结果。

得到效率工资以后,便可通过人均工资收入估算人均人力资本存量 $h$ 。同样将三种方法得到的结果列于表3进行对比。

表3 人均人力资本存量计算结果对比

城市	1992	1996	2000	2004	2008	2009	2010
北京	3.4686 (11.4280) [23.9]	4.1775 (12.8670) [30.1]	3.9811 (7.5632) [34.4]	4.9671 (10.0730) [68.8]	10.0790 (36.8452)	11.4933 (46.6324)	12.8425 (53.9342)
天津	3.8114 (16.7800) [22.5]	4.5302 (19.1370) [21.7]	4.3091 (11.2370) [25.6]	4.9475 (9.7689) [36.4]	6.1830 (10.2849)	6.1417 (8.9143)	6.2020 (7.3694)
上海	8.2557 (60.3560) [30.4]	8.1313 (47.3900)	7.6496 (31.7260) [49.4]	8.5670 (34.8790) [71.0]	12.8509 (48.3197)	14.3735 (56.0938)	15.2119 (59.3691)

资料来源:根据作者整理数据计算得到。

从表3同样可以看到三个特点:第一,朱平芳和徐大丰(2007)的结果呈现不断增长的趋势,符合中国经济发展的实际情况,因此该方法对城市人力资本的估算具有一定的合理性;第二,当使用朱平芳和徐大丰(2007)的方法,对本研究使用的省际数据进行估算时,得到的结果在1992年到2004年间却呈现出下降趋势,这与中国经济发展的实际情况存在明显的出入,即使从人力资本存量的角度看下降趋势依然存在(见图1);第三,用本研究调整后的方法和省际数据得到的估算结果表现出人均人力资本不断增长的趋势,比较符合中国实际,具有一定合理性。

在人均人力资本估算结果的基础上,将其与在岗职工数相乘就可以得到省际人力资本存量的估算结果。表4中列出的是我国部分年份各省(市、区)人力资本存量的估算结果,其中括号内是人均人力资本存量<sup>①</sup>。

从表4中所列的估算结果可以看到,从1978年到2010年间,我国东部、中部和西部,不论是人力资本存量还是人均人力资本存量都存在逐渐增长的趋势。其中东部增长速度最快,中部次之,西部最差。这一点也与中国经济增长和人力资本存量积累的基本情况相符,进一步说明该估算方法与结果具有合理性。

<sup>①</sup> 由于篇幅限制,表4中只包含了主要年份数据,有兴趣的读者可向作者索要全部数据。

表 4 各省（市、区）主要年份人力资本存量估算结果

省(市、区)	1978	1992	2007	2008	2009	2010
北 京	1082.9 (3.8265)	1653.1 (3.4686)	3677.5 (7.6793)	5123.3 (10.0790)	7118.4 (11.4933)	8304.4 (12.8425)
天 津	485.0 (2.2349)	1120.6 (3.8114)	1081.1 (6.1413)	1101.2 (6.1830)	1238.5 (6.1417)	1275.5 (6.2020)
河 北	1074.0 (2.4135)	2420.2 (3.5208)	1549.9 (3.1876)	1565.9 (3.3270)	1907.7 (3.7923)	1992.7 (3.8352)
山 西	559.1 (2.0862)	1557.0 (3.3679)	1260.0 (3.4357)	1303.6 (3.5647)	1463.3 (3.7929)	1534.2 (3.8898)
内 蒙 古	331.7 (2.5714)	1464.5 (3.7293)	687.6 (2.8232)	661.6 (2.7457)	661.6 (2.6915)	676.2 (2.7134)
辽 宁	4717.1 (7.3020)	10026.6 (9.5930)	3502.6 (7.4058)	3082.2 (6.3460)	3206.2 (6.2922)	3042.5 (5.8718)
吉 林	1464.6 (4.1491)	2361.3 (4.3607)	898.8 (3.5038)	804.1 (3.1431)	819.4 (3.0889)	770.2 (2.8781)
黑 龙 江	2704.9 (4.3279)	2707.9 (3.0761)	2196.7 (4.7980)	1972.2 (4.6575)	2453.7 (5.2312)	2198.7 (4.7799)
上 海	1610.9 (3.8172)	4220.3 (8.2557)	3478.3 (11.8658)	3924.2 (12.8509)	5539.6 (14.3735)	5976.3 (15.2119)
江 苏	1431.9 (2.4645)	2327.2 (2.5740)	2290.3 (3.4324)	2366.5 (3.5412)	2783.8 (3.8592)	3007.5 (3.9378)
浙 江	671.3 (2.1446)	1377.8 (2.8037)	3070.5 (4.8531)	3552.3 (5.1531)	4975.6 (6.1131)	5553.0 (6.2846)
安 徽	602.6 (1.9501)	1217.4 (2.3983)	1344.5 (4.1555)	1288.5 (3.9948)	1509.6 (4.1949)	1548.2 (4.1514)
福 建	605.2 (2.9377)	1286.4 (3.7746)	2242.6 (5.2239)	2265.0 (5.1294)	2478.8 (5.2299)	2708.1 (5.3400)
江 西	517.6 (1.9312)	1165.5 (2.8538)	859.5 (3.1259)	859.6 (3.1241)	948.5 (3.2749)	993.0 (3.3389)
山 东	811.9 (1.7274)	1684.2 (2.0103)	2777.5 (3.1575)	2804.6 (3.2139)	3353.0 (3.6357)	3673.2 (3.8415)
河 南	582.8 (1.3876)	1342.7 (1.7994)	1926.6 (2.7562)	1903.8 (2.7515)	2229.9 (3.0349)	2205.3 (2.9339)
湖 北	1126.9 (2.4498)	3183.4 (4.3495)	1741.8 (3.9251)	1638.8 (3.6996)	2066.6 (4.2429)	2281.0 (4.4702)
湖 南	617.1 (1.6954)	1571.3 (2.7105)	1468.2 (3.5812)	1486.7 (3.5283)	1745.9 (3.6296)	1769.6 (3.4992)

续表

省(市、区)	1978	1992	2007	2008	2009	2010
广东	1165.6 (1.9557)	2670.6 (3.1123)	3999.8 (4.0775)	4203.3 (4.2619)	4963.1 (4.7042)	5561.3 (4.9720)
广西	466.7 (1.9609)	1117.0 (3.3444)	910.2 (3.3487)	922.3 (3.3886)	1082.0 (3.5898)	1032.0 (3.2590)
四川	1045.8 (1.6040)	3075.9 (3.1304)	1847.2 (3.5498)	1950.2 (3.6871)	2400.5 (4.2533)	2537.3 (4.4469)
贵州	238.3 (1.4102)	473.9 (2.0096)	750.9 (3.6181)	717.1 (3.5899)	843.4 (3.8566)	876.3 (3.9064)
云南	794.3 (3.6774)	1716.8 (5.5724)	1409.4 (5.0208)	1377.1 (4.8030)	1557.3 (4.9909)	1454.0 (4.5048)
陕西	480.9 (1.8713)	986.3 (2.4994)	1259.9 (3.8081)	1296.5 (3.9044)	1433.5 (4.0680)	1428.0 (3.9145)
甘肃	211.4 (1.8713)	968.2 (3.8420)	936.6 (4.9163)	921.9 (4.8985)	1077.0 (5.5807)	1005.3 (5.1740)
青海	134.0 (2.8518)	206.8 (3.0496)	139.1 (3.2292)	152.9 (3.4308)	180.0 (3.5559)	177.2 (3.3688)
宁夏	114.8 (2.6081)	368.2 (5.1209)	308.3 (5.4735)	267.3 (4.8877)	242.6 (4.1757)	223.4 (3.7658)
新疆	978.4 (4.3103)	979.6 (3.0775)	817.0 (3.4120)	837.8 (3.5030)	1056.8 (4.2370)	1026.2 (4.0249)

注：括号中标注的是人均人力资本存量。

资料来源：根据作者整理数据计算得到。

为了更加直观地反映我国省际人力资本的变化情况，并更为直观地对比本研究修正后的方法与朱平芳和徐大丰（2007）方法的估算结果，我们通过折线图来反映四个省份人力资本的变化趋势，具体如图1所示。图1中描绘了使用两种不同估算方法得到的北京、广东、山西和广西四个省（市、区）从1978年到2010年的省际人力资本存量。需要说明的是，考虑到人力资本估算结果是无量纲的，所以两种方法得到的结果没有绝对意义，只有相对意义。因此，我们将两种方法得到的结果分列两图。其中左图为本研究修正后的估算结果，右图为使用朱平芳和徐大丰（2007）的估算方法和省际数据得到的结果。

从左图中可以清晰地看到人力资本的持续增长，而且东部地区北京与广东的增长速度明显高于中部与西部地区的山西与广西，这些结果都比较符合中国的实际情况。当然，在左图中也很容易发现，四个省份的估算结果在1998年附近均出现一定的下降趋

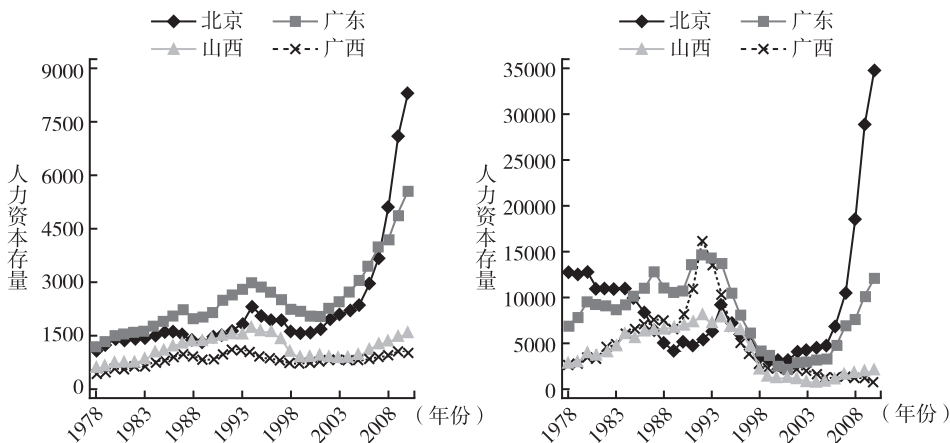


图 1 不同方法下四个省份人力资本存量估算结果对比

注：右为朱平芳和徐大丰（2007）的方法，左为调整后结果。

资料来源：根据作者整理数据计算得到。

势。这主要是因为相关数据的统计口径发生了变化，1998 年以前统计的是职工数和职工平均工资，而 1998 年以后统计的是在岗职工数和在岗职工平均工资。

作为对比，右图中的省际人力资本数据明显存在着先降低后增加的问题，而且山西与广西 2010 年的人力资本存量水平甚至比 1978 年还要低，这显然有悖于实际。因此，从估算结果与中国实际情况对比的角度出发，本研究修正后的估算结果更为合理，在一定程度上也是对省际人力资本存量更为准确的反映。

## 五 计量检验

本文调整后得到的估计结果是否合理还需要经过进一步的考察，下面利用面板数据的计量经济模型进行检验。根据式（1），通过对方程两边同时取对数可得如下模型：

$$\log Y = \beta \log A + \alpha \log K + \beta \log H + \gamma \log Z \quad (11)$$

其中， $Y$  代表总产出， $K$  代表物质资本存量， $H$  代表人力资本存量， $Z$  代表综合制度变量。由于人力资本是不断上升的，技术进步的趋势主要由人力资本决定，故可以将劳动有效性  $A$  视为常量，并将公式（11）中的  $\beta \log A$  写成常数项。

至此，进行计量分析的难点就集中在了选择综合制度变量  $Z$  上。由于式（11）在

模型形式设定上的限制,变量  $Z$  必须是连续的。此外,考虑到中国经济制度变革具有地域和时序差别,不同省份的制度情况必然影响该地区人力资本积累速度。因此,综合制度变量  $Z$  还必须体现出地区差异。由于极少有将人力资本与制度变迁相联系的文献,而人力资本积累又与经济增长有着密不可分的关系,因此可以使用对经济增长影响较大的制度变量来衡量  $Z$ 。Zhang et al. (2001) 发现国有企业与经济增长之间有着密切联系。借鉴他们的研究成果,我们将国有经济的情况作为综合制度变量进行考察。综合以上分析,构建如下回归模型:

$$\log GDP_{it} = \_const + \beta_1 \log K_{it} + \beta_2 \log H_{it} + \beta_3 \log SEPA_{it} + a_i + \varepsilon_{it} \quad (12)$$

该模型中,  $i = 1, 2, \dots, 28$  表示不同的省(市、区)<sup>①</sup>,  $t = 1978, 1979, \dots, 2005$  表示不同的年份。 $\log GDP$  是以 1952 年为基年的实际 GDP 的对数,  $\log K$  表示以 1952 年为基年的实际社会固定资本存量的对数,  $\log H$  表示人力资本存量的对数,  $\log SEPA$  表示国有工业企业总产值占全部工业企业总产值比重的对数,  $\_const$  为常数项,  $a_i$  表示不随时间改变的个体因素,  $\varepsilon_{it}$  为随机干扰项,满足白噪声的假设。

回归方程中共有四个解释变量,具体的数据来源如下。其中,前三个变量的数据比较容易获取,  $\log GDP$  主要通过历年《中国统计年鉴》中的名义 GDP 与 GDP 指数核算而得,  $\log K$  的数据主要来自张军等(2004)的估算结果,  $\log H$  则是本研究的估算结果。

最后一个制度变量  $\log SEPA$  的来源相对复杂。该指标主要涉及两个统计数据,一是国有工业企业总产值,二是全部工业企业总产值。其中,国有工业企业总产值有比较一致的统计数据,1998 年及以前的数据主要来自《新中国 50 年统计资料汇编》,1998 年以后的数据主要来自历年的《中国区域经济统计年鉴》和历年的分省统计年鉴。但是,全部工业企业总产值的统计口径有过较大的改变,1998 年以后国家统计局不再统计规模以下工业企业总产值。因此,为了统计口径的一致,还需要获取规模以下工业企业的总产值。具体的获取途径主要有四种:其一,《新中国 55 年统计资料汇编》中统计了部分地区 1999 年到 2004 年的全部工业企业总产值,比如广东和福建等;其二,各省历年统计年鉴中有部分地区汇报了全部工业总产值,比如广西和甘肃等;其三,各地区报刊杂志等相关报道,比如吉林等;其四,各地区的国民经济和社会发展统计公报中也有少量相关数据。当然,在此基础上还有一些缺失数据,这些数据主要利用全部工业企业增加值进行估算。由于具体的估算方法并非本文的重点,在此不

① 不包括西藏、重庆和海南。

做详细介绍。通过多种渠道和多种估算方法补充后，我们将该指标收集到 2005 年。由于只是检验本研究人力资本估算结果的有效性，因此，我们认为 1978 年到 2005 年的样本并不会对检验的效果产生显著的不利影响。

因为样本数据共包含 28 个个体的 28 期观测数据，时序相对较长，因此可能存在一定的时序特征。为了避免伪回归的问题，确保面板数据中的时序是平稳序列，需要先对面板数据进行单位根检验。目前来看，常见的检验面板数据单位根的统计量有不少，比如 LLC (Levin et al., 2002)、IPS (Im et al., 2003)、MADF (multivariate augmented Dickey-Fuller) (Sarno & Taylor, 1998; Taylor & Sarno, 1998) 和 Fisher-PP (Maddala & Wu, 1999)。不同的统计量在检验上具有不同的特点，以及不同的适用范围。其中，MADF 检验基于似无相关模型 (SURE)，因而要求面板数据的时序数大于截面个体数，因此本文的数据并不符合要求。Fisher-PP 以单个个体的 ADF (augmented Dickey-Fuller) 检验为基础来构造  $p$  值，其优点主要体现在可用于非平衡面板。LLC 基于固定效应估计，假设所有序列均服从同一个一阶自回归 AR(1) 过程，不仅允许个体固定效应的存在，而且考虑了截面异质性和干扰项的序列相关。与 LLC 不同，IPS 放宽了相同 AR(1) 的假设，即 LLC 是相同单位根检验，IPS 是不同单位根检验。由于接下来主要使用固定效应估计模型，因此，本文使用 LLC 和 IPS 两个统计量检验是否存在单位根 (面板单位根检验通常包括相同单位根和不同单位根检验两个内容)，具体结果如表 5 所示。

从表 5 的检验结果可以看到，LLC 检验在不包含趋势项时显著拒绝了存在单位根的原假设，说明这些变量不存在相同的单位根。IPS 检验在包含趋势项时大都显著拒绝了存在单位根的原假设，即也不存在不相同的单位根。其实，现有统计量的渐近性假设普遍要求时序足够长。因此，这种时序与截面相当的面板数据不用过于关注单位根和伪回归的问题，该检验结果也仅供参考。

表 5 面板数据单位根 LLC 与 IPS 检验结果

解释变量	LLC		IPS	
	不含趋势项	含趋势项	不含趋势项	含趋势项
实际 GDP 对数	-5.581 **	-9.879	0.490	-2.479 **
社会固定资本存量对数	-9.049 ***	-13.914 ***	0.853	-2.445 **
人力资本存量对数	-2.867 ***	-8.425	-2.072 ***	-1.932
国有工业企业占比对数	-1.439 *	-9.908	-13.371 ***	-2.728 ***

注：\*  $p < 0.10$ ，\*\*  $p < 0.05$ ，\*\*\*  $p < 0.01$ 。

资料来源：根据作者整理数据计算得到。

单位根检验只是避免了伪回归的问题,但是,面板数据还可能存在着序列相关、异方差和截面相关等问题。为了修正这些问题,我们分别选择基于 Driscoll & Kraay (1998) 稳健性标准误的固定效应方法和 PCSE (panel-corrected standard errors) 方法对模型进行估计。Driscoll & Kraay (1998) 标准误是在固定效应的估计结果中,在获取标准误时考虑了序列相关、异方差和截面相关的问题。PCSE 是采用最小二乘法 (OLS) 或 Prais-Winsten 估计方法,但在计算标准误时考虑了序列相关、异方差和截面相关等问题,而且该方法也比较适合截面不大,时序不长的方块面板,因此比较适合本研究中 28 个截面和 28 期时序的数据结构。具体的估计结果如表 6 所示。

表 6 省际人力资本存量结果检验

解释变量	(1)	(2)	(3)	(4)	(5)
社会固定资本 存量对数	0.7306 *** (0.0255)	0.5377 *** (0.0137)	0.7505 *** (0.0204)	0.7489 *** (0.0129)	0.7489 *** (0.0183)
人力资本存量对数	0.3121 *** (0.0399)	0.5143 *** (0.0105)	0.2939 *** (0.0260)	0.3595 *** (0.0157)	0.3595 *** (0.0238)
国有工业企业 占比对数	-0.1588 *** (0.0293)	-0.3880 *** (0.0185)	-0.0958 *** (0.0193)	-0.1092 *** (0.0143)	-0.1092 *** (0.0191)
常数项	-0.5172 *** (0.1654)	0.1132 (0.1262)	-0.7432 *** (0.1728)	-1.1428 *** (0.1259)	-1.1428 *** (0.1515)
样本数	784	784	784	784	784
调整的 R <sup>2</sup>	—	0.9695	0.9616	0.9922	0.9922

注:第(1)列,使用 Driscoll & Kraay (1998) 稳健性标准误的固定效应估计;第(2)列,考虑截面相关性与截面异方差的 PCSE 估计;第(3)列,考虑截面相关性、截面异方差和截面具有相同序列相关系数的 PCSE 估计;第(4)列,考虑截面异方差与截面间具有不同序列相关系数的 PCSE 估计;第(5)列,考虑截面相关性、截面异方差和截面间具有不同序列相关系数的 PCSE 估计;\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ , 括号内为标准误。

资料来源:根据作者整理数据计算得到。

从五个估计结果看,第二列的估计系数,即不考虑单个的截面序列相关问题时的估计系数与其他各列之间存在比较显著的差异。这说明,在有 28 期时序的面板数据中,考虑单个截面的序列相关性是必要的。从其他四个估计结果看,“社会固定资本存量对数”的估计结果基本保持在 0.74 左右,“国有工业企业占比对数”的估计结果在 -0.1588 到 -0.0958 之间,变化不大。而我们的重点检验对象“人力资本存量对数”,其估计结果在 0.2939 到 0.3595 之间,也相对稳定。其中,最后一列的估计结果,即同时考虑截面相关性、截面异方差和截面间具有不同的序列相关性时,人力资本弹性的估计系数为 0.3595,与我们前面的假设也最为接近,与章上峰和许冰(2009)的估计



结果也较为接近。不仅如此，该估计结果也不能拒绝三个估计系数之和等于 1 的原假设，这与前面设置的规模报酬不变的生产函数模型也是一致的。所以，我们认为同时考虑截面相关性、截面异方差和截面间具有不同的序列相关系数的模型设置最为合理。同时也说明，我们得到人力资本的估计方法具有一定的合理性与适用性。

## 六 结论

人力资本存量是影响经济增长的重要变量。由于对其量化存在很多困难，因此一直是学者研究的焦点。而省际人力资本存量的估算由于其特殊性，又具有一些难点。比如，省际数据的获取存在困难，省际之间的人才流动影响地区间的人力资本存量。虽然朱平芳和徐大丰（2007）提出了城市人力资本存量估算的可行方法，但如果将其直接应用在省际数据中会出现问题。因此，本研究在朱平芳和徐大丰（2007）提出的城市人力资本存量估算方法的基础上，调整了生产函数的基本形式与基本假设，并对物质资本弹性与人力资本弹性的设置做出了相应调整。最后，利用修正方法得到的估算结果不论从数值的角度看，还是从统计分析的角度看，都比较符合中国的实际情况。因此，我们认为由该方法估算得到的数值可以用于进一步的实证研究。

最后需要指出的是，本研究的估算框架仍然存在进一步改进的空间。首先，模型中劳动有效性  $A$  与综合制度变量  $Z$  同人力资本之间的关系可以考虑更为复杂的非线性关系。其次，两个重要的参数  $\alpha$  和  $\beta$  的设定方式也可以进行深入的分类型讨论。最后，综合制度变量  $Z$  的代理变量或衡量方式，也有许多值得讨论的内容。

### 参考文献：

- 蔡昉、都阳（2003），《“文化大革命”对物质资本和人力资本的破坏》，《经济学（季刊）》第 4 期，第 795 - 806 页。
- 高慧、余利霞（2007），《四川省与全国人力资本对经济增长的作用比较分析》，《理论与改革》第 6 期，第 89 - 91 页。
- 廖楚晖（2006），《中国人力资本和物质资本的结构及政府教育投入》，《中国社会科学》第 1 期，第 23 - 33 页。
- 张帆（2000），《中国的物质资本和人力资本估算》，《经济研究》第 8 期，第 65 - 71 页。

- 张军、吴桂英、张吉鹏 (2004), 《中国省际物质资本存量估算: 1952 - 2000》, 《经济研究》第 10 期, 第 35 - 44 页。
- 章上峰、许冰 (2009), 《时变弹性生产函数与全要素生产率》, 《经济学 (季刊)》第 2 期, 第 551 - 568 页。
- 周兴、陈东娇 (2007), 《广东经济增长中的人力资本贡献》, 《江苏商论》第 8 期, 第 147 - 148 页。
- 朱平芳、徐大丰 (2007), 《中国城市人力资本的估算》, 《经济研究》第 9 期, 第 84 - 95 页。
- Barro, Robert J. & Jong-Wa Lee (1993). International Comparisons of Educational Attainment. *NBER Working Paper*, No. 4349.
- Driscoll, John C. & Aart C. Kraay (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics*, 80(4), 549 - 560.
- Im, Kyung So, M. Hashem Pesaran & Yongcheol Shin (2003). Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics*, 115(1), 53 - 74.
- Levin, Andrew, Chien-Fu Lin & Chia-Shang James Chu (2002). Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of Econometrics*, 108(1), 1 - 24.
- Lucas, Robert Jr. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(1), 3 - 42.
- Maddala, G. S. & Shaowen Wu (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics*, 61 (S1), 631 - 652.
- Mankiw, N. Gregory, David Romer & David N. Weil (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, 107(2), 407 - 437.
- Mulligan, Casey B. & Xavier Sala-i-Martin (1995). A Labor-Income-Based Measure of the Value of Human Capital: An Application to the States of the United States. *NBER Working Paper*, No. 5018.
- Romer, David (1996). *Advanced Macroeconomics*. New York: McGraw-Hills Companies.
- Romer, Paul M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5), S71 - S102.
- Sarno, Lucio & Mark P. Taylor (1998). Real Exchange Rates under the Recent Float: Unequivocal Evidence of Mean Reversion. *Economics Letters*, 60(2), 131 - 137.

- Taylor, Mark P. & Lucio Sarno (1998). The Behavior of Real Exchange Rates during the Post-Bretton Woods Period. *Journal of International Economics*, 46(2), 281 – 312.
- Zhang, Anming, Yimin Zhang & Ronald Zhao (2001). Impact of Ownership and Competition on the Productivity of Chinese Enterprises. *Journal of Comparative Economics*, 29(2), 327 – 346.

## **The Estimation of China's Provincial Human Capital Stock 1978 – 2010: Improvement of Existing Methods**

Luo Zhi<sup>1,2</sup> & Zhao Anping<sup>3</sup>

(School of Social Sciences, Tsinghua University<sup>1</sup>;

Institute of Management, Beijing Academy of Social Sciences<sup>2</sup>;

Beijing Office, China Banking Regulatory Commission<sup>3</sup>)

**Abstract:** The estimation of provincial human capital stock is essential to the macroeconomic analysis of China using panel data methods. Zhu Pingfang and Xu Dafeng (2007) proposed a method in estimating city-level human capital stock. However, when this method is used in estimating provincial human capital stock, we find the results do not conform to China's actual conditions. We argue the direct reason is that the elasticity of physical capital and human capital could not reflect the reality of China. Thus, through adjusting the production function and relaxing basic assumptions, we modify the elasticity of human capital and physical capital. At last, through econometric test, we find the estimation using this revised method is in accord with the reality of China. So this estimation can be used in other empirical researches.

**Keywords:** human capital stock, average wage, panel data

**JEL Classification:** C23, E24, J24

(责任编辑：贾 朋)