

## 中国的收入机会不平等

——基于2013年中国家庭收入调查数据的研究

董丽霞\*

---

**内容提要** 中国当前的收入不平等问题饱受质疑。影响收入不平等的因素中,既有自身不能控制的环境因素,也有个人努力因素。根据Roemer(1998)的研究,如果人们的收入与环境因素无关则意味着机会平等;反之,如果收入与环境因素相关,则意味着机会不平等。本文基于机会公平理论框架,将影响收入的变量分为环境和努力两类,通过随机占优方法检验机会不平等的存在性,并用参数方法衡量中国的收入机会不平等,分析结果不平等中机会不平等所占的比例。结果发现:2013年中国收入机会不平等指数约为0.10,环境造成的机会不平等约占结果不平等的1/5;夏普里分解得知,家庭收入和出生地是造成收入不平等的最主要因素;分位回归结果表明,中等收入群体的收入机会更加不平等。

**关键词** 收入机会不平等 努力 环境 随机占优 夏普里分解

---

### 一 引言

近年来,中国的收入不平等问题饱受质疑。Xie & Zhou(2014)使用中国家庭追踪调查、中国综合社会调查、中国劳动力动态调查等五组数据的研究结果一致表明,中国的家庭收入基尼系数在0.53~0.55之间,且在最近三十年急剧升高,从1980年的

---

\* 董丽霞,中国社会科学院经济研究所,电子邮箱:donglixia@163.com。感谢国家自然科学基金青年项目(批准号:71603029)和北京市社会科学基金青年项目(编号:14JYC028)的资助。

0.3 上升到 2010 年的 0.53。根据北京师范大学李实教授团队的最新研究，如果包括高收入人群样本，从 2007 到 2013 年，中国的基尼系数从 0.49 扩大到 0.55<sup>①</sup>。中国的收入不平等不仅高于大多数发展中国家，甚至高于美国等发达国家。这种收入不平等会使代际流动固化，影响长期经济增长和社会稳定，还可能影响人们的价值观念、社会规范和政治文化等。因此有必要采取措施扭转这一局面。收入不平等既有个人努力因素造成的不平等，也有个人无法控制的外在因素（如父母教育程度、家庭收入和父母职业等家庭背景）造成的机会不平等。

收入不等的研究在中国有很多。这些研究从财政分权、对外贸易、人口特征、腐败、金融结构、户口和职业隔离等不同层面揭示了收入不平等产生的原因（陈刚，2011；杨俊、王佳，2012；张宗益等，2012；马万里等，2013；吴晓刚、张卓妮，2014；韩军等，2015）。但现有文献大多关注的是“收入”这一结果的平等，而很少有文献从机会平等角度探讨剔除了努力因素后的“机会”平等问题。由于缺乏理论指导，现有关于中国收入不平等的实证研究极少计算考虑了个人努力因素后的机会不平等，更鲜有文献将中国的收入不平等分解为个人努力造成的不平等和环境造成的机会不平等。其原因主要在于努力程度无法度量，因此机会不平等难以衡量，实证研究难以入手。

陈宗胜和李清彬（2011）在分析再分配倾向选择的公平信念动机时提到，如果一个人的收入更多是由出生地点、父母的经济社会地位、关系等不可控因素决定，而不是努力奋斗的结果，则人们对再分配的支持倾向就较大。这意味着，人们并不排斥个人努力造成的不平等，而是不愿忍受个人无法控制的外在因素造成的不平等。因此，更合理的公平目标应该是机会公平，应该考虑个人努力因素。在个人努力相同的情况下，个人的收入等结果变量与其家庭背景等外在因素无关，这就是机会公平。

Roemer（1998）提出机会公平理论（Equality of opportunity，简称 EOP）的完整分析框架，强调个人努力造成的结果差异的合理性。他将不平等的产生原因分为两类：一类是个人无法控制的外在因素，即 Roemer（1998）所说的“环境”造成的不平等；另一类是个人的努力因素造成的不平等。相对来说，人们对个人努力造成的结果差异接受程度较高。Roemer（1998）的框架成为识别和衡量机会不平等的理论基础。随后，

<sup>①</sup> 参见李实教授在 2016 中国经济发展论坛上的演讲，<http://finance.sina.com.cn/meeting/2016-11-19/doc-ifxxwrwk1463456.shtml>。

一批文献利用这一框架对机会不平等的衡量、估计方法和实证应用等进行了研究。Checchi & Peragine (2010) 利用机会公平框架衡量了意大利的收入机会不平等,发现2000年机会不平等在意大利总的收入不平等中所占比例约为20%。Ferreira & Gignoux (2011) 则用2000年前后的拉丁美洲数据进行了研究,发现这些国家机会不平等占总的收入不平等的比重为30%左右。

目前,很少有学者将机会平等理论应用于中国收入机会不平等问题的研究。中国收入机会不平等程度有多大?结果不平等中有多少来自机会不平等,多少来自努力因而可以视作合理的不平等?这些问题还没有一个令人信服的答案。本文力图回答上述问题。在理论上,本文基于Roemer (1998) 的机会平等分析框架,衡量剔除努力因素后的收入机会不平等,并将总的收入不平等分解为机会不平等和努力程度差异导致的不平等。在实证上,本文分别使用随机占优和参数方法进行实证分析。本文首先使用一阶随机占优方法检验机会不平等的存在性;然后使用基于计量回归的参数方法考察父母教育水平、家庭收入、父母职业和出生地等多个环境因素对收入的影响;在此基础上使用反事实分析方法和不平等指数,计算剔除努力因素后中国的收入机会不平等程度并估计机会不平等在结果不平等中所占的比例;最后用夏普里分解方法寻找对机会不平等影响最大的环境因素,从而为政策制定提供依据。

本文的结构安排如下:第二部分是分析框架;第三部分使用随机占优方法初步检验收入机会不平等的存在性;第四部分在机会公平理论框架基础上进行实证检验,衡量收入机会不平等程度和估计各环境因素的贡献度;最后是结论和启示。

## 二 分析框架

### (一) 机会平等的内涵与理论基础

传统的不平等研究主要关注的是收入、教育、健康等结果的不平等。机会公平理论与传统公平理论的区别在于其考虑了个人努力因素,并将努力造成的不平等视为合理的不平等。

政治哲学家Rawls (1958) 最先提出机会公平的概念。他最早将个人责任作为道德上合理的不公平而引入公平理论,但主要强调的还是结果公平。直到Rawls (1971), 公平理论的研究视角才开始从结果公平转向机会公平,其理念认为公平不是要求结果的平等,而是机会的平等。Rawls (1971) 之后,Sen (1980)、Dworkin (1981)、

Arneson (1989) 和 Cohen (1989) 围绕公平的内涵和目标进行了讨论。Sen (1980) 强调个人可用的功能品集合的平等。Dworkin (1981) 讨论了资源公平，资源包括个人不应负责的生来就有的身体生理环境。Arneson (1989) 认为社会的目标应该是福利机会平等。Cohen (1989) 认为分界点不应该在偏好（或选择）和资源之间，如果个人偏好弱势环境中形成，个人不应为偏好负责。上述文献从哲学上提出机会不平等的思想，并开始区分道德上可接受的和不可接受的不平等。但由于个人负责的部分即努力程度难以量化，机会不平等的衡量缺乏科学的工具，而仅仅停留在理论层面的探讨上，机会公平理论无法应用于实践。

20 世纪 90 年代以来，经济学家开始关注个人责任，并将研究视角从结果不平等转向机会公平问题。其中，Roemer (1998) 开创性地提出一个机会平等的完整分析框架（如图 1 所示）。他将影响结果的因素分为环境和努力两类。前者是个人不能控制的因素，由所有的社会、遗传因素组成，如种族、性别、出生地和社会经济出身等家庭背景；后者是个人可以控制的因素。值得说明的是，这里的努力是广义的努力，不仅包括个人可以选择的学习时间长短、学习的努力程度或努力时间等狭义的个人努力程度，还包括环境中没有包括的天赋等因素（这样的定义高估了努力造成的不平等，低估了环境引起的不平等）。相应地，总的结果不平等可以分解为不合理的（由超出个人控制的环境因素引起）和合理的不平等（由个人可以控制的努力因素引起）。个人仅应该为其自身选择的努力因素负责，不应该为其无法控制的环境因素负责。因此，源自外生环境的不合理的不平等在道德上是不可接受的，应该被补偿；而源自努力的合理的不平等在道德上是可以接受的，不需要任何干预。Roemer (1998) 将“机会平等”定义为结果与环境无关。一个“公平社会”的目标是机会平等，而非最终结果的平等。公平的原则不是要求每个人获得的最终结果平等，而是使人们获得某一结果的机会平均分布，给每个人提供相同的机会，使个人最终获得的成果取决于其自身的选择。促进机会平等的政策目标是缩小面临不同机会的人（处于不同环境但努力程度相同的人）获得成就（如收入、成绩等）或得到优势资源（如上大学）的差距。机会公平理论强调个人对其最终结果的影响，这与传统强调结果公平的平均主义思想截然不同。

## （二）机会不平等的度量方法

衡量机会不平等程度的难点在于剥离努力程度。Roemer (1998) 的框架成为识别和衡量机会不平等的理论基础。随后，Peragine (2004)、Fleurbaey & Maniquet (2011) 等一批文献对机会不平等的衡量、估计方法、实证应用和政策影响等都进行了研究。

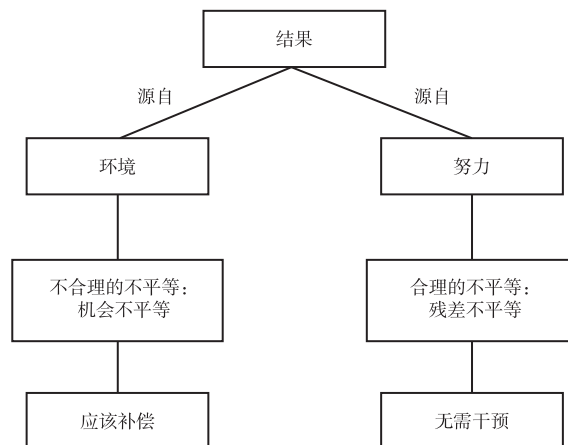


图 1 机会平等理论框架

衡量机会不平等的经济模型大体上分为两类。一类是假定每个人有给定的机会集，衡量机会不平等就是比较个人机会集的联合分布。这种方法直接考察机会，因而被称为“直接方法”。但在实际中，机会往往观察不到，这一方法很难在实证中得到应用。另一类是“间接方法”。这一方法不关注人们的机会分布，而是看给定机会的结果分布。间接方法满足两个基本原则，即补偿原则和回报原则。源自环境的不平等，超出个人的控制范围，是不合理的，需要补偿，这是补偿原则。这一原则要求社会补偿由超出个人控制的因素所导致的结果差异。源自努力差异的结果不平等，道德上是合理的，努力得到相应回报是合理的，这是回报原则。回报原则要求社会尊重个人努力所得到的回报，而不应补偿努力所造成的结果差异。一个追求结果完全平等的社会，满足补偿原则，但违反回报原则；一个完全自由放任的社会，满足回报原则，但违反补偿原则。而机会平等的社会则要求同时满足补偿原则和回报原则。补偿原则有事前补偿和事后补偿两种，主要区别在于补偿先于努力还是看到努力水平之后再补偿。如果补偿先于努力，补偿原则要求无论什么环境的人，其机会集都相同。如果补偿在努力之后，补偿原则要求无论处于什么环境，努力程度相同的人其结果均相同。这意味着，事前补偿的目的是消除不同环境之间机会集的不平等，将最终的不平等视为努力造成的结果差异；事后补偿的目的是消除努力程度相同群体内部的不平等，因此这种补偿要求事前知道个人的努力程度。但在实际中很难准确度量个人的努力程度。回报原则强调相同环境下努力程度不同造成的结果差异的合理性。因此，事前补偿与回报原则

一致，而事后补偿与回报原则未必一致。并且，事前补偿原则避免了努力程度难以度量的问题。因此，为了避免回报原则需要度量努力程度，本文实证部分主要基于事前补偿方法来衡量机会不平等。

事前补偿方法主要考察的是努力程度相同但环境不同的人之间的结果不平等。如果努力程度相同但环境不同的人结果相同，则不存在机会不平等。基于这个思路，衡量机会不平等的问题变成了考察将实际收入用剥离了努力后的收入代替的反事实收入的不平等。

基于事前补偿原则度量机会不平等程度需要两个步骤：第一步是构建反事实分布，剥离掉努力以及合理的不平等，考察反事实分布反映的不合理的不平等；第二步是衡量机会不平等。文献中主要有两种方法估计反事实分布以及不同环境分布之间的差距。第一种方法是序数方法，即随机占优方法。第二种方法是首先进行计量回归然后使用不平等指数的两步法。接下来的实证部分将分别使用这两种方法来考察中国收入的机会不平等问题。本文先使用随机占优方法比较不同环境下的收入分布差异，并据此检验机会不平等的存在性；再基于计量回归模型，计算反事实收入的不平等指数，即机会不平等指数。

### 三 检验收入机会不平等的存在性

#### （一）数据和变量

本文使用的数据是北京师范大学中国收入分配研究院联合国内外专家共同完成的中国家庭收入调查（CHIP）数据。本文使用的是2013年的城镇样本数据。这一数据包含子女的收入以及个人特征、家庭背景等信息。CHIP 2013样本有5115个家庭，经过匹配后父母和子女信息完整的家庭有2470个。

本文结果变量使用的是子女在2013年的年工资收入。环境变量则包括了影响子女收入但其个人无法控制的多个因素，包括家庭背景和个人特征变量。具体地，环境分为两类：第一类变量是家庭背景变量，包括父母教育程度、父母职业、父母收入等；第二类是子女的个人特征，包括性别、出生地、民族、年龄等。

#### （二）随机占优

在本文中，结果是收入，机会集指的是给定环境 $s$ 后收入 $x$ 的分布，即 $F(x|s)$ 。机会公平意味着不同环境下人们的收入分布相同。这样，衡量机会不平等就是考察不同环境下收入分布函数的差异程度。因此，第一步就是通过考察不同环境下结果的条

件分布是否有差异来检验机会不平等的存在性。这就是 Lefranc et al.(2008) 所做的,使用随机占优概念和统计检验来比较不同环境之间的机会分布。其中,环境根据父母教育、父母职业和父母收入等家庭背景来划分。

随机占优方法最初多用于金融学文献中,如 Rothschild & Stiglitz (1970)。后被 Van de Gaer (1993) 尝试用于比较不同环境下的机会集。随机占优方法有一阶随机占优和二阶随机占优之分。其中一阶随机占优主要考虑的是回报问题,二阶随机占优则同时考虑了回报和风险问题。为了降低复杂性,本文不考虑个人的风险偏好,只使用一阶随机占优方法。如果影响个人收入的因素是确定的,使用一阶随机占优方法就足够了。根据 Lefranc et al.(2008),一阶随机占优的定义如下:

环境  $s_1$  一阶随机占优 (first-order stochastic dominance, 简称 FSD) 于环境  $s_2$ , 即  $s_1 \geq_{FSD} s_2$ , 当且仅当对于任何收入  $x$ , 这两个环境下的收入条件分布函数的关系是  $F(x | s_1) \leq F(x | s_2)$ 。这意味着对于任何给定分位数量的人群,分布  $s_1$  上的平均收入总是高于分布  $s_2$  上的平均收入;对于任何一档收入,分布  $s_1$  比之分布  $s_2$ , 低于该收入的人的比例更少。即不考虑风险偏好的情况下,每个人都会更偏好环境  $s_1$ 。这样,就可以得出结论,存在收入机会不平等。如果条件分布相同,即  $F(x | s_1) = F(x | s_2)$ , 环境对于收入分布没有影响,则意味着机会公平。

即使无法观察到所有的环境,这个方法也依然有效。Lefranc et al.(2008) 证明,即使没有描述全部环境,在给定环境条件下的分布相同也是机会平等的必要条件。这个方法提供了一个用非参数方法检验收入机会不平等的参考。

由于这里研究的问题是收入机会平等,结果变量是收入水平。从理论上来说,影响个人收入的环境因素主要有父母受教育水平、家庭收入状况、父母的职业等。根据随机占优的思想,如果不同家庭背景的人其收入既不重合也不交叉,说明存在收入的机会不平等。这些不同环境下人们的收入分布之间的距离反映了不同环境在影响收入方面的优势。如果不同环境下人们的收入分布较为接近,说明机会不平等程度较小。具体结果如图 2 至图 4 所示。

图 2 显示的是根据父母教育程度分类的长子(女)年收入的累积分布函数。由此可见,父母受教育程度越高,其子女的收入水平越高。父母受教育程度为大学及以上的子女收入普遍最高,高于父母受教育程度为中学的子女,父母受教育程度为小学及以下的子女收入最低。因此,父母受教育程度较高的子女收入明显占优于父母受教育程度较低的子女。这个结果表明,父母受教育程度不同的人之间存在显著的机会不平等。也就是说,在父亲或母亲受教育程度代表的与个人努力无关的环境条件下,存在

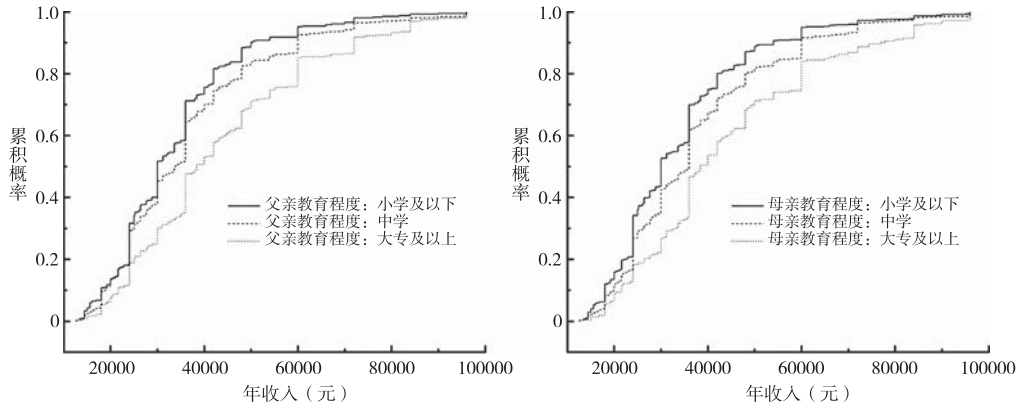


图2 基于父母教育程度分类的子女收入累积分布函数

资料来源：根据 CHIP 2013 数据计算得到。

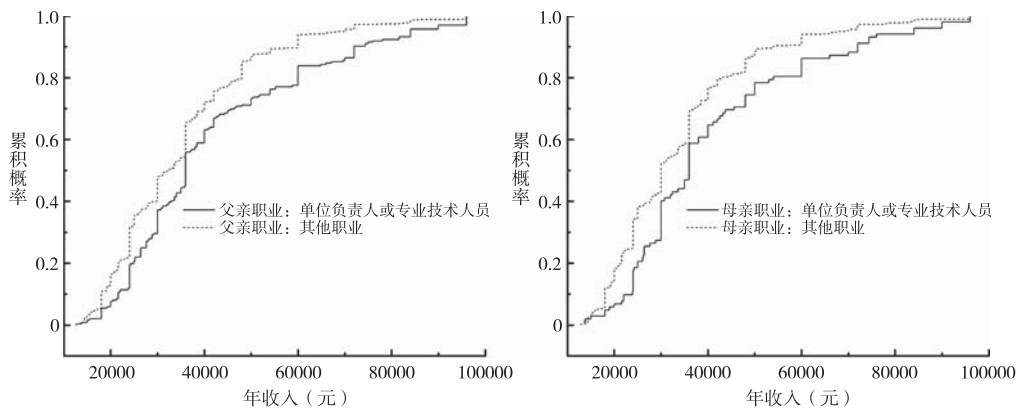


图3 基于父母职业分类的子女收入累积分布函数

资料来源：根据 CHIP 2013 数据计算得到。

明显的收入机会不平等现象。

图3显示的是父母不同职业条件下的长子(女)年收入分布。由图可见,父母职业为国家机关、党群组织、企事业单位负责人或专业技术人员的子女收入分布显著优于父母职业为办事人员、商业服务业人员、农林牧渔和水利业生产人员、生产、运输设备操作人员及其他从业人员等其他职业子女。父母职业身份的高低意味着获得资源的能力高低。这个结果表明,父母职业身份较高的子女收入分布明显占优于父母



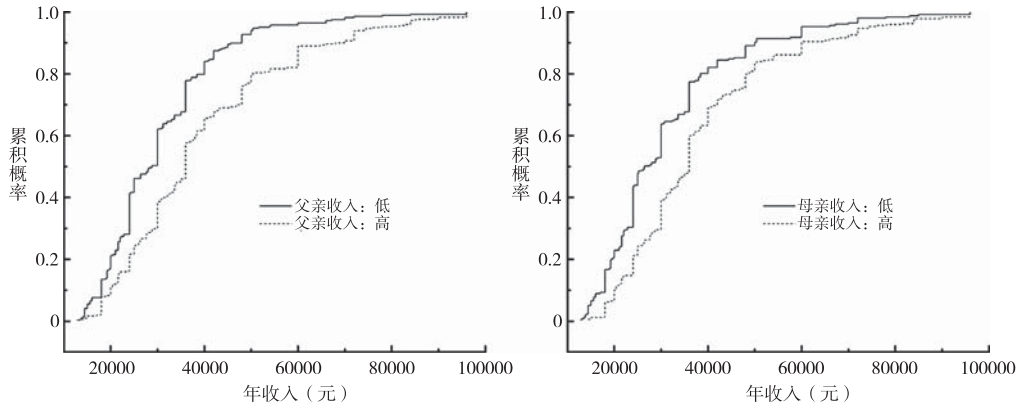


图4 基于父母收入水平分类的子女收入累积分布函数

资料来源：根据CHIP 2013数据计算得到。

职业身份较低子女。这说明，在父母职业身份代表的与个人努力无关的环境条件下，存在明显的收入机会不平等现象。

图4显示的是父母不同收入类型条件下的长子(女)年收入分布。由于随机占优考察的是不同环境条件下的收入分布，环境变量需是离散型变量，而收入是连续变量，这里将父母收入以中位数为界，分成高收入、低收入两组。由图可见，父母收入高的子女收入分布显著占优于父母收入低的子女。这说明，在父母收入代表的与个人努力无关的环境条件下，存在明显的收入机会不平等现象。

## 四 实证分析

### (一) 基准模型

由于本文关心的是衡量机会不平等程度，而非研究机会平等的政策效应，相关性足以满足研究需要，无需建立因果关系模型。因此，模型中只考虑与结果相关的环境和努力变量，不再考虑其他控制变量。估计系数也不能解释为因果关系，而是反映了环境对结果的影响程度，即机会不平等。

为了定量研究家庭背景不同的子女的收入机会平等程度，家庭背景对子女收入的影响可以用如下结构方程(structural model)估计：

$$Y_{it} = \alpha C_{it} + \beta E_{it} + u_{it} \quad (1)$$

$$E_{ii} = \kappa C_i + v_{ii} \quad (2)$$

其中， $Y_{ii}$ 是结果变量即子女的收入， $C_i$ 是环境变量如父母教育程度， $E_{ii}$ 是努力程度， $u_{ii}$ 、 $v_{ii}$ 是误差项， $\alpha$ 、 $\beta$ 、 $\kappa$ 是估计系数。方程（1）反映的是环境对子女收入的直接影响，方程（2）反映的是环境通过影响努力而对结果的间接影响。由于这里只是估计机会不平等，衡量观察到的环境对结果变量的影响，不需要估计上述结构方程和推导因果关系，只需要估计诱导方程（reduced-form equation）和相关关系。将方程（2）代入方程（1）中，可得如下诱导方程：

$$Y_{ii} = \delta C_i + \varepsilon_{ii} \quad (3)$$

对数线性化后，可以得到如下方程：

$$\ln Y_{ii} = \gamma C_i + \sigma_{ii} \quad (4)$$

其中， $Y_{ii}$ 是结果变量即子女的年工资收入， $C_i$ 是环境变量向量， $\sigma_{ii}$ 是误差项， $\gamma$ 是估计系数。可以通过普通最小二乘（OLS）方法来估计诱导方程，推导出方程中被环境解释的部分。

由于方程（4）是诱导方程，环境通过影响努力进而对结果变量的任何影响都反映在回归系数 $\hat{\gamma}$ 中。因此，方程（4）反映了环境对结果的全部影响，包括直接影响和通过影响努力而影响结果的间接影响。参照 Ferreira & Gignoux（2011）的研究，构建收入的反事实分布的参数估计如下：

$$\bar{Y} = \exp[\hat{\gamma} C_i] \quad (5)$$

$\bar{Y}$ 是反事实的收入分布。 $\hat{\gamma}$ 是方程（4）OLS回归估计出的系数， $\hat{\gamma} C_i$ 反映了环境对子女收入的全部影响，即机会不平等。如果用预测值代替每个子女的收入，所有相同环境的子女应该有相同的收入。因此，在机会完全平等的社会中，所有子女的预测收入都相同。而这些反事实分布的不平等只是源自环境差异，与努力无关。因此，机会不平等可以通过反事实收入的不平等来衡量。

基准回归结果如表1所示。基准模型（1）中使用的环境变量是子女性别、子女年龄、子女民族、父母受教育程度、出生地，模型（2）和模型（3）将父母受教育程度分别替换为父母职业和父母收入。这些模型的结果显示，女性子代收入显著偏低，子女年龄对收入的影响显著为正，子女出生在东部地区对其收入有显著正效应，父母教育、父母收入和父母职业都对子女收入有显著影响。这说明，收入机会不平等的事实存在。模型（4）将父母受教育程度、父母收入和父母职业全部包括后，父母职业和父母教育变量不再显著，原因可能在于这两个变量与父母收入直接相关。考虑到子女自

身的教育和职业对其收入有着直接的影响,模型(5)增加了子女的教育,模型(6)又进一步增加了子女职业。这两个模型考察了环境变量影响子女收入的渠道。这两个模型中,父母职业和父母教育变量也不显著,这反映了这两个变量通过影响子女的职业和教育而影响子女的工资收入。因此,基准模型的结论是,父母教育程度、父母职业和家庭收入等家庭背景和性别、出生地等环境变量对子女收入有着显著的影响。这表明,中国存在收入机会不平等问题。

表1 基准回归

		模型(1)	模型(2)	模型(3)	模型(4)	模型(5)	模型(6)
子女性别		-0.0793 ***	-0.0821 *	-0.0731 ***	-0.0916 **	-0.1110 ***	-0.0838 **
子女年龄		0.0053 ***	0.0162 ***	-0.0013	0.0218 ***	0.0207 ***	0.0214 ***
子女民族		0.1320 ***	-0.0085	0.1526 ***	-0.0361	-0.0418	-0.0687
出生地	中部地区	-0.0343	-0.0703	-0.0169	-0.0899	-0.0692	-0.0787
	东部地区	0.1624 ***	0.0991 *	0.1689 ***	0.0636	0.0689	0.0534
父亲收入		—	—	0.1178 ***	0.1914 ***	0.1746 ***	0.1616 ***
母亲收入		—	—	0.1346 ***	0.1863 ***	0.1514 ***	0.1546 ***
父亲职业		—	0.1760 ***	—	0.0893	0.0701	0.0607
母亲职业		—	0.0868	—	0.0358	0.0387	-0.0046
父亲教育	中学	0.0102	—	—	-0.0453	-0.0530	-0.0597
	大专及以上	0.1597 ***	—	—	0.0138	-0.0540	-0.0621
母亲教育	中学	0.1076 ***	—	—	0.0527	-0.0065	0.0145
	大专及以上	0.1621 ***	—	—	-0.0054	-0.0936	-0.0593
子女教育	初中	—	—	—	—	0.3163	0.3075
	高中	—	—	—	—	0.3758 *	0.3510
	职高、技校、中专	—	—	—	—	0.4161 *	0.3960 *
	大专	—	—	—	—	0.4422 **	0.4273 **
	大学本科	—	—	—	—	0.5394 **	0.5137 **
	研究生	—	—	—	—	0.8212 ***	0.7840 ***

续表

		模型(1)	模型(2)	模型(3)	模型(4)	模型(5)	模型(6)
子女职业	国家机关、党群组织、企事业单位负责人	—	—	—	—	—	0.3914 **
	专业技术人员	—	—	—	—	—	0.3669 **
	办事人员和有关人员	—	—	—	—	—	0.1701
	商业、服务业人员	—	—	—	—	—	0.2063
	生产、运输设备操作人员及有关人员	—	—	—	—	—	0.3447 **
	军人	—	—	—	—	—	0.2343
	不便分类的其他从业人员	—	—	—	—	—	0.2102
常数项	9.9542 ***	9.8566 ***	10.0686 ***	9.6410 ***	9.3131 ***	9.0782 ***	
样本数	1833	459	1936	459	458	457	
R <sup>2</sup>	0.0983	0.1030	0.0884	0.1916	0.2334	0.2684	

注：“子女性别”变量的参照组为女性；“子女民族”变量的参照组为汉族；“出生地”变量的参照组为西部地区；“父亲教育”变量的参照组为小学及以下；“母亲教育”变量的参照组为小学及以下；“子女教育”变量的参照组为小学及以下；“子女职业”变量的参照组为农林牧渔和水利业生产人员。为节省空间，表中没有汇报标准误。显著性水平：\*\*\* 1%，\*\* 5%，\* 10%。

资料来源：根据 CHIP 2013 数据计算得到。

## （二）估计收入机会不平等指数

通过回归方法，可以证实中国存在收入机会不平等。家庭背景等个人不能控制的环境变量对子女的收入具有显著影响。但这种机会不平等的程度到底有多大，还需要进一步计算。本文将在回归模型（4）基础上，计算反事实的收入  $\tilde{Y}$ 。由于不平等指数中，只有广义熵指数满足对称性、转移性、等价不变性、总体复制性以及可加可分性等性质，因此这里使用广义熵指数计算  $I(\tilde{Y})$  即机会不平等的绝对水平（IOA），在此基础上，除以子女实际收入的广义熵指数可以计算机会不平等的相对水平（IOR）。即：

$$IOA = I(\tilde{Y}) \quad (6)$$

$$IOR = \frac{I(\tilde{Y})}{I(Y)} \quad (7)$$

广义熵指数系列中有多个不平等指标，例如泰尔指数即 GE(1)，变异系数的平方即 GE(2)，均值对数离差（mean logarithmic deviation，简称 MLD）指数即 GE(0)。

为了计算总的、努力造成的不平等和剔除努力后环境造成的不平等，需要利用一个可分解的不平等指数，使得总的不平等等于努力引起的不平等与环境造成的不平等即机会不平等之和。均值对数离差（MLD）指数满足这一要求。其计算公式为：

$$MLD(X) = \frac{1}{N} \sum_{i=1}^N \ln \frac{\mu}{\chi_i} \quad (8)$$

其中， $N$  为目标群体的子女数， $\chi_i$  为目标群体中每个子女的收入， $\mu$  为这些子女的平均收入。将这几个不平等指数都进行估计，结果如表 2 所示。从表 2 可以发现，这几个不平等指数得出的结果差异不大。不论是绝对水平，还是相对水平，在剔除自身努力程度后，中国都存在收入机会不平等问题。根据 MLD 指数结果，中国总的收入不平等指数约为 0.10，其中，环境造成的机会不平等为 0.02。环境造成的机会不平等占总的收入不平等的比例约为 21.11%。

表 2 2013 年中国城镇收入机会不平等指数

不平等指标	结果不平等	机会不平等	机会不平等/结果不平等
GE(0)	0.1040	0.0220	0.2111
GE(1)	0.1035	0.0222	0.2140
GE(2)	0.1126	0.0228	0.2026

资料来源：根据 CHIP 2013 数据计算得到。

### （三）贡献度分析

本部分基于夏普里分解方法分解各种环境因素对收入机会不平等的贡献。夏普里值（Shapley Value）的概念来自博弈论，是合作博弈中的核心概念。它刻画了局中人在合作博弈中的地位。基于回归的夏普里值分解方法由 Shorrocks（2013）提出<sup>①</sup>，逐渐成熟并被许多学者用于不平等指标的分解。这种运用夏普里值基于回归方程对不平等指标进行分解的方法，不仅可以考察不平等主要由哪些因素决定，还可以量化这些因素对不平等的贡献。与传统的分解方法相比，夏普里分解方法有两个主要优点：分解结果没有残差，分配给每个解释变量的影响之和等于

① 该文的早期版本是 1999 年的一篇工作论文，即：Shorrocks, Anthony (1999). Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value. University of Essex and Institute for Fiscal Studies。

观察到的被分解变量的变化；分解方法适用于任何模型、任何函数形式、任何不平等指标的分解，在分解时可以使用任意数量与类型的因素和交互作用项甚至代理变量。对这一方法的更详细介绍可参见 Shorrocks (2013)、Wan (2004) 等文献。

本部分将使用夏普里分解方法，在基准模型(4)回归结果的基础上，分析各环境变量对收入不平等的贡献度。为了避免扭曲收入变量的分布及含义，因变量用的是收入本身，而不是模型(4)中收入的对数。表3是最终的分解结果。其中，家庭收入变量分别使用父亲收入、家庭总收入、人均家庭收入和等价家庭收入(令等价家庭规模分别为0.5和0.75)来考察<sup>①</sup>。表4则显示了用家庭消费即家庭总消费、人均家庭消费和等价家庭消费(令等价家庭规模分别为0.5和0.75)代替家庭收入的分解结果。

表3 环境变量(收入)对子女收入机会不平等的贡献：基于回归模型的夏普里值分解

环境	父亲收入		家庭总收入		人均家庭收入		等价家庭收入 (a=0.5)		等价家庭收入 (a=0.75)	
	夏普里值	%	夏普里值	%	夏普里值	%	夏普里值	%	夏普里值	%
子女性别	0.0039	1.93	0.0035	1.66	0.0036	1.58	0.0035	1.57	0.0036	1.57
子女年龄	0.0322	15.95	0.0286	13.62	0.0261	11.40	0.0276	12.25	0.0269	11.75
出生地	0.0773	38.21	0.0734	34.88	0.0691	30.18	0.0695	30.83	0.0689	30.15
父亲职业	0.0161	7.94	0.0171	8.13	0.0159	6.93	0.0160	7.10	0.0158	6.93
母亲教育	0.0130	6.42	0.0123	5.84	0.0107	4.66	0.0110	4.89	0.0108	4.71
家庭收入	0.0597	29.54	0.0755	35.87	0.1035	45.24	0.0977	43.35	0.1026	44.89
总体	0.2021	100	0.2103	100	0.2289	100	0.2254	100	0.2285	100

资料来源：根据 CHIP 2013 数据计算得到。

<sup>①</sup> 每个家庭的规模和年龄结构未必相同，因此家庭的福利水平不能只看家庭收入水平。不平等文献中通常用等价家庭规模来调整以得到可比的等价家庭收入。参见 Hindriks & Myles (2006)。

表 4 环境变量（消费）对子女收入机会不平等的贡献：基于回归模型的夏普里值分解

环境	家庭总消费		人均家庭消费		等价家庭消费 ( $\alpha = 0.5$ )		等价家庭消费 ( $\alpha = 0.75$ )	
	夏普里值	%	夏普里值	%	夏普里值	%	夏普里值	%
子女性别	0.0041	2.08	0.0045	2.18	0.0043	2.13	0.0044	2.16
子女年龄	0.0278	14.28	0.0258	12.56	0.0269	13.25	0.0263	12.86
出生地	0.0776	39.81	0.0752	36.69	0.0753	37.08	0.0750	36.63
父亲职业	0.0180	9.26	0.0168	8.18	0.0171	8.41	0.0168	8.23
母亲教育	0.0138	7.07	0.0119	5.79	0.0124	6.12	0.0121	5.89
家庭收入	0.0536	27.50	0.0709	34.60	0.0670	33.01	0.0701	34.25
总体	0.1948	100	0.2050	100	0.2030	100	0.2048	100

资料来源：根据 CHIP 2013 数据计算得到。

从表 3 和表 4 中可以有以下发现。家庭收入和出生地是造成子女收入机会不平等的最主要环境变量。从不同收入变量的结果来看，家庭收入对收入差异的贡献率最高为 44.89%；出生地略低，但其贡献率仍高达 39.81%。父亲职业、母亲教育程度和年龄对子女收入机会不平等也有重要影响。父亲职业对收入机会不平等的贡献率为 9.26%，母亲教育程度的贡献率为 7.07%。年龄对收入机会不平等的贡献率为 15.95%。性别对收入机会不平等的影响贡献率为 2.18%，这是贡献率最低的环境变量。

以上结果说明，在影响收入机会公平的因素中，家庭环境方面最迫切需要解决的问题是家庭收入、出生地和父亲职业、母亲受教育程度引起的不公平。因此，在政策层面，为促进机会公平，应该补偿这些环境因素引起的机会不公，促进机会公平。

#### （四）分位回归模型

进一步地，还可以研究家庭背景的影响是否随子女的收入分布而变化。这种家庭背景影响的异质性可以通过 Koenker & Bassett (1978) 提出的分位回归 (quantile regression) 方法来进行估计。分位回归估计子女收入分布的不同分位上家庭背景对子女收入的影响。如果分位回归估计结果在各收入分位上相对一致，说明家庭背景对子女收入的影响在不同收入群体之间不存在显著的异质性。如果各收入分位上的分位回归估计结果显著不同，则说明家庭背景对于子女收入的影响在不同收入群体之间存在显著的异质性。图 5 所示的是父亲年收入的数值代表的家庭背景对长子（女）年收入

人对数值的分位回归和普通最小二乘回归结果的系数值。由图 5 可以看出，家庭背景对不同分位子女收入的影响呈现“倒 U”型。收入最低和最高的人群，家庭背景对其影响较小；而中等收入人群，家庭背景对其影响较大。因此，比之收入最低和最高的人群，中等收入群体的收入机会更加不公平。这可能是目前许多中产阶级家庭陷入焦虑的原因。他们拼命奋斗，力图为子女在教育、健康等方面创造好的环境，在子女进入学校和社会竞争之前，中产阶级家庭之间的竞争就已经展开。

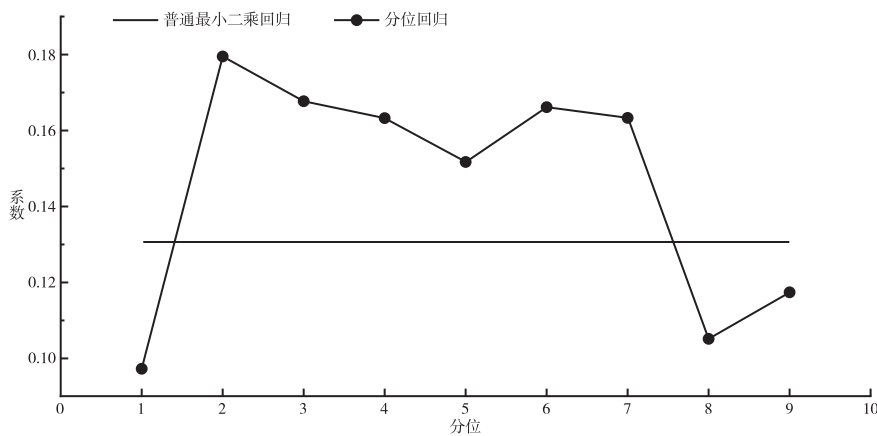


图 5 家庭背景对子女收入影响的分位回归结果

资料来源：根据 CHIP 2013 数据计算得到。

## 五 结论和启示

本文基于 Roemer (1998) 的机会平等分析框架，分析了剔除努力因素后中国的收入机会不平等问题。将影响结果的变量分为环境和努力两类，通过随机占优方法检验了机会不平等的存在性，并基于回归方法衡量剥离了努力因素的收入机会不平等问题。在此基础上，使用广义熵指数计算了剔除努力因素影响的收入机会不平等以及结果不平等中机会不平等所占的比例。本文的研究结果说明，中国确实存在收入机会不公平问题，父母教育水平、家庭收入、父母职业和出生地区等多个环境因素对子女的收入结果均有显著影响。父母受教育程度越高，父母工作越好，家庭收入越高，子女的收入越高。具体发现如下：中国收入机会不平等指数约为 0.10，结果不平等中约有 21.11% 来自机会不平等，环境造成的机会不平等约占结果不平等的 1/5。基于回归的



夏普里分解得知,家庭收入和出生地是造成收入不平等的最主要因素。父亲职业、母亲教育程度和年龄对收入机会不平等也有重要影响。而性别对收入机会不平等的影响贡献率最低。分位回归结果表明,比之收入最低和最高的人群,中等收入群体的收入机会更加不平等。

由于机会不平等会进一步恶化收入不平等,因此,有必要通过政策的改变来促使收入机会更加公平。这意味着,未来政府在制定再分配政策时要充分考虑个人的努力程度。此外,在此基础上,对于家庭收入较低、出生于落后地区的群体,政府要通过再分配等政策给予支持,最大程度补偿环境差异,减少不公平的机会,从而降低机会不公平。

值得一提的是,本文对中国收入机会不平等的估计偏低。由于不可能观察到构成个人环境的所有特征(如内在天赋或能力),因此本文估计的机会不平等只是环境造成的不平等即机会不平等的下限值。相应地,为弥补机会不平等所需的再分配需求偏小。因此,政府在进行再分配政策时力度应该更大一些。未来如果获得更好的数据,可以用更好的方法估计机会不平等的整个区间,并对中国不同区域进行比较研究,从而可以得到关于中国收入机会不平等的更全面更准确更切合国情的结论。

## 参考文献:

- 陈刚(2011),《腐败与收入不平等——来自中国的经验证据》,《南开经济研究》第5期,第113-131页。
- 陈宗胜、李清彬(2011),《再分配倾向决定框架模型及经验验证》,《经济社会体制比较》第4期,第35-46页。
- 韩军、刘润娟、张俊森(2015),《对外开放对中国收入分配的影响——“南方谈话”和“入世”后效果的实证检验》,《中国社会科学》第2期,第24-40页。
- 马万里、李齐云、张晓雯(2013),《收入分配差距的财政分权因素:一个分析框架》,《经济学家》第4期,第13-23页。
- 吴晓刚、张卓妮(2014),《户口、职业隔离与中国城镇的收入不平等》,《中国社会科学》第6期,第118-140页。
- 杨俊、王佳(2012),《金融结构与收入不平等:渠道和证据——基于中国省际非平稳异质面板数据的研究》,《金融研究》第1期,第116-128页。
- 张宗益、杜鹏、汪锋(2012),《人口特征对收入分配的影响研究——基于深圳和重庆

- 数据的分析》，《人口研究》第7期，第78-90页。
- Arneson, Richard (1989). Equality and Equal Opportunity for Welfare. *Philosophical Studies*, 56(1), 77-93.
- Cecchi, Daniele & Vito Peragine (2010). Inequality of Opportunity in Italy. *Journal of Economic Inequality*, 8(4), 429-450.
- Cohen, Gerald (1989). On the Currency of Egalitarian Justice. *Ethics*, 99(4), 906-944.
- Dworkin, Ronald (1981). What is Equality? Part 2: Equality of Resources. *Philosophy & Public Affairs*, 10(4), 283-345.
- Ferreira, Francisco & Jérémie Gignoux (2011). The Measurement of Inequality of Opportunity: Theory and an Application to Latin America. *Review of Income and Wealth*, 57(4), 622-657.
- Fleurbaey, Marc & François Maniquet (2011). *A Theory of Fairness and Social Welfare*. Cambridge: Cambridge University Press.
- Hindriks, Jean & Gareth Myles (2006). *Intermediate Public Economics*. Cambridge: The MIT Press.
- Koenker, Roger & Gilbert Bassett Jr (1978). Regression Quantiles. *Econometrica*, 46(1), 33-50.
- Lefranc, Arnaud, Nicolas Pistolesi & Alain Trannoy (2008). Inequality of Opportunities vs. Inequality of Outcomes: Are Western Societies All Alike? *Review of Income and Wealth*, 54(4), 513-546.
- Peragine, Vito (2004). Measuring and Implementing Equality of Opportunity for Income. *Social Choice and Welfare*, 22(1), 187-210.
- Rawls, John (1958). Justice as Fairness. *The Philosophical Review*, 67(2), 164-194.
- Rawls, John (1971). *A Theory of Justice*. Cambridge: Harvard University Press.
- Roemer, John (1998). *Equality of Opportunity*. Cambridge: Harvard University Press.
- Rothschild, Michael & Joseph Stiglitz (1970). Increasing Risk: I. A Definition. *Journal of Economic Theory*, 2(3), 225-243.
- Sen, Amartya (1980). Equality of What? In Sterling McMurrin (ed.), *The Tanner Lectures on Human Values (Volume 1)*. Cambridge: Cambridge University Press, pp. 197-220.
- Shorrocks, Anthony (2013). Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value. *Journal of Economic Inequality*, 11(1), 99-126.

- Van de Gaer, Dirk (1993). Equality of Opportunity and Investment in Human Capital. Ph. D. Dissertation, Katholieke Universiteit Leuven.
- Wan, Guanghua (2004). Accounting for Income Inequality in Rural China: A Regression-based Approach. *Journal of Comparative Economics*, 32(2), 348 – 363.
- Xie, Yu & Xiang Zhou (2014). Income Inequality in Today's China. *Proceedings of the National Academy of Sciences*, 111(19), 6928 – 6933.

## **Inequality of Income Opportunity in China: Evidence from the 2013 CHIP Data**

Dong Lixia

(Institute of Economics, Chinese Academy of Social Sciences)

**Abstract:** Income inequality has become a serious social problem in China. Two groups of factors are found to affect inequality. One includes ‘circumstances’ factors, in which individuals are embedded and generally have little control; and the other includes ‘personal’ factors which individual could change through personal efforts. According to Roemer (1998), equality of opportunity implies outcomes are uncorrelated with social circumstances, while a correlation between the two suggests inequality of opportunities. Under this theoretical framework, this paper tests the existence of inequality of income opportunity with a stochastic dominance method and estimates the inequality level using parametric methods. The result shows that the index of inequality of income opportunity was 0.10 in 2013, and about 1/5 of inequality of outcomes could be attributed to inequality of opportunity in China. A further Shapley decomposition suggests that family income and place of birth are two most important factors leading to inequality in opportunities. Results of quantile regression shows that the inequality of income opportunity is more serious among the middle-income group.

**Keywords:** inequality of income opportunity, efforts, circumstances, stochastic dominance, Shapley decomposition

**JEL Classification:** C46, D31, J71

(责任编辑: 西 贝)