

中国流动人口收入不平等中的机会不平等测度

——基于事前估计视角

刘 林 李 猛*

内容提要 本文基于中国流动人口动态监测调查数据,测度了流动人口中的机会不平等程度,并对其影响因素进行 Shapley 分解。结果表明,从 2010 年到 2017 年,流动人口的收入不平等指数和机会不平等指数均呈上升趋势,农村户籍和高年龄组群体的机会不平等更为严重,且性别是造成机会不平等的最主要因素。在考虑到环境对机会不平等的间接效应后,流动人口机会不平等的相对值在 22.38%~24.61% 之间。最后,本文通过反事实模型对机会不平等间接效应做了进一步的稳健性检验。

关键词 流动人口 机会不平等 反事实模型

一 引言

改革开放四十年来,中国经济社会发展取得了举世瞩目的成就,在摆脱贫困和增进人民福祉方面实现了历史性巨大进步。四十年里,中国也经历了世界上最大规模的国内劳动力流动,流动人口成为中国经济高速增长和城镇化进程的重要推动力量(呼倩等,2019;叶文平等,2018)。而在经济高速增长的同时,收入差距也在不断扩大,收入不平等现象愈发突出,引起社会各界广泛关注和讨论。平衡收入分配日益成为中

* 刘林,石河子大学经济与管理学院,电子邮箱:aspiration83lin@sina.com;李猛(通讯作者),石河子大学经济与管理学院,电子邮箱:limon0816@163.com。作者感谢国家自然科学基金项目“新疆南疆四地州的农户可持续生计:生计变迁与生计调适”(批准号:71963029)、“兵团英才”项目对本文研究的资助。

国跨越中等收入陷阱的关键所在（马占利、邹薇，2018）。

联合国经济和社会事务部发布的《2020年世界社会报告》指出，目前，无论是发展中国家还是发达国家，不平等水平都处于历史最高位，这一严峻形势将会加剧分歧，减缓社会经济发展步伐。不平等因素是社会矛盾的重要根源，如何构建兼顾效率与公平的收入分配体系，是政府必须优先考虑的问题（李莹、吕光明，2019）。

目前，国内学术界关于不平等的研究更多集中在对结果不平等的讨论。然而，收入分配领域中的结果不平等并不能如实反映产生收入差距的深层次原因（李莹，2019），且并非所有的收入不平等都是不公平、不合理的。区分道德上可接受和不可接受的不平等是最近四十年来哲学平均主义思想最重要的贡献（Roemer & Trannoy, 2016）。从上个世纪七八十年代开始，以罗尔斯为代表的西方政治哲学家开始提出一种关于平等主义的新方法，将个人责任加入到对平等与道德的讨论中。此后，平等主义理论的发展可以被描述为用机会平等来取代结果平等的趋势。Roemer（1993）在总结前人研究的基础上开创性地提出机会均等的“环境（circumstances）-努力（efforts）”二元分析框架。他将“机会均等”本质上定义为一种重要的结果，称之为“优势（advantages）”，独立于环境分配的情况，从而也正式确立了与哲学的区别。

在Roemer（1993）关于机会均等的定义中，把那些个体能够控制且能为之负责的因素称为“努力”。值得注意的是，此处的努力是广义的努力，不仅包括个体工作的努力程度、学习的努力程度及职业的选择，还包括个人的天赋、运气等因素（董丽霞，2018）。那些个体不能掌控也不能为之负责的因素称为“环境”，比如个人的性别、种族、出生地、父母的受教育水平和家庭的社会经济地位等。个体的结果（收入、财富或健康）是在努力变量和环境变量共同作用下产生的（Betancort et al., 2019）。因此，Roemer（1998）将总体的结果不平等分成由努力因素导致的不平等即“努力不平等”和由环境因素导致的不平等即“机会不平等”两个部分。不平等在道德上令人反感的部分来自那些个人无法通过努力改变并影响结果的因素（Cecchi & Peragine, 2010）。

国外学者对机会不平等做了大量研究，积累了丰富的研究文献。这些研究主要集中在两个方面：一方面是在收入分配领域，研究和测算各个国家在不同阶段的机会不平等程度（Cecchi & Peragine, 2010；Bourguignon et al., 2007；Ferreira & Gignoux, 2011）；另一方面是在卫生健康领域，探讨环境和努力对健康不平等的影响与贡献（Trannoy et al., 2010；Dias, 2009；Davillas & Jones, 2020）。总结这些研究文献，我们发现关于机会不平等的研究主要有以下几个特点：

首先，从估计机会不平等的角度来看，文献中区分了事前（Ex-ante）估计法和事

后 (Ex-post) 估计法 (Fleurbaey & Peragine, 2013)。两种方法的不同主要在于对机会均等的定义 (Aaberge et al., 2011)。事前法的重点是各个机会集之间的不平等 (Fleurbaey et al., 2017)。事前法将总体划分为不同的类别 (type), 即具有相同环境的个体的集合。事后法则关注相同努力水平下个人之间的结果不平等。事后法将总体划分为不同的群组 (tranche), 即具有相同努力水平的个体的集合, 衡量机会不平等就是衡量群组内的结果不平等。当所有人付出相同程度的努力而达到相同的结果时, 机会均等就得以实现 (Juárez & Soloaga, 2014)。在实际的研究中, 事前法比事后法更为常见, 原因在于努力变量难以直接观测且难以准确度量。

其次, 从机会不平等的测量方法上来看, 文献区分了“参数法”和“非参数法” (Ferreira & Gignoux, 2011)。参数法估计需要设定包含环境变量和努力变量的收入决定方程, 利用估计系数和可观测的环境变量的均值, 模拟出消除环境影响后的“反事实”收入分布, 得到的反事实收入与计算得到的实际收入之间的差距, 即是收入分配的机会不平等。非参数法则不需要设定具体的收入决定方程, 而是依据环境变量或努力变量对样本对象进行分组, 得到的组间不平等即为机会不平等。相对于参数法, 非参数法的不足之处在于, 当环境变量过多时分组数会成倍增加, 而测算不平等需要类别内部有足够多的样本, 因此就需要庞大的数据量来做支撑 (雷欣等, 2018)。所以, 在测度机会不平等时, 事前视角的参数法的使用更为普遍。在国内的研究中, 李莹和吕光明 (2018)、刘波等 (2015, 2020)、江求川等 (2014)、史新杰等 (2018) 的研究都采用参数估计法。然而, 也有少数学者使用非参数法或基于事后视角开展研究, 如龚峰等 (2017) 基于事后估计视角, 利用非参数法测算了努力对机会不平等的影响, 董丽霞 (2018) 利用非参数法下的一阶随机占优检验了机会不平等的存在性, 同时使用参数法测算了剔除努力影响后的机会不平等程度。

最后, 总结目前关于机会不平等的既有研究, 我们发现, 众多学者从理论、估计角度和测量方法上提出了诸多讨论和改进策略, 使得机会不平等的测量更加规范和具体。然而, 在目前的测量中仍有两点需要关注。一是未观测到的环境变量导致的机会不平等的低估。在实际测量中并非所有的环境变量都可被观测到, 未观测到的环境变量可能就错误地归于努力和运气 (Roemer & Trannoy, 2016), 这就导致基于事前视角的估计都是对机会不平等的下限估计 (Niehues & Peichl, 2014; Brunori et al., 2019)。二是努力与环境的相关性。本文在上文介绍事后估计时提到了努力因素存在难以观测和准确度量的问题。在现有文献研究中, 努力变量常取个人的受教育程度 (Roemer et al., 2016)。然而, 将其视为个人努力变量存在争议, 因为初等和中等教育是在个体的

儿童和青少年时期进行的，更多受家庭情况和父母的影响。因此，越来越多的学者在测量机会不平等时不仅关注环境因素对结果的直接影响，还关注环境通过努力因素对结果的间接影响（Palomino et al., 2019）。

目前，国内学者关于机会均等的研究不论是在理论上还是在研究范围上仍需进一步拓展。鲜有学者将机会均等理论应用到流动人口群体进行研究。流动人口中收入的机会不平等程度有多大？收入的不平等中由环境因素所导致的机会不平等占多大的比重？哪些环境因素是机会不平等生成的主要因素？本文力图回答上述问题。本文依托 Roemer（1993, 1998）提出的机会均等的“环境-努力”分析框架，充分考虑环境因素与努力因素的相关性，基于事前估计视角和参数法考察可观测到的性别、户籍、出生地等环境变量对收入的影响，并使用 Shapley 分解得出各个环境因素对机会不平等的贡献度。在此基础上，本文逐步引入受环境因素影响的努力变量，来测算环境通过努力因素对机会不平等的间接效应，最后采用反事实模型，对上述关于环境对机会不平等的间接效应做进一步的稳健性检验。

本文的结构安排如下：第二部分是机会不平等测度模型的构建；第三部分是变量的选取和数据的描述性统计；第四部分是实证结果与分析，在计量回归的基础上，测算机会不平等的绝对量和相对量；第五部分是环境对机会不平等间接效应的检验；最后是结论与启示。

二 机会不平等的测度

（一）机会均等的内涵和基准模型的构建

根据 Roemer（1993）关于机会均等的定义，当个体的收入不受其环境的影响时，机会均等就能实现，如下所示：

$$F^t(y) = F^m(y), \forall t, m \quad (1)$$

其中， $F^t(y)$ 表示类别 t 中的收入分配， $F^m(y)$ 表示类别 m 中的收入分配。式（1）表明，只要不同个体付出相同的努力，就能得到相同的收入分配，而环境于结果无关紧要，也就实现了机会均等。相反，如果一种分配占主导，机会均等就不再成立。

收入的不平等有两种来源，分别是环境（circumstances）和努力（efforts）（加上运气和未观察到的随机项）。因此，个体 i 的收入 y_i 可表示为：

$$y_i = f(c_i, e_i, \mu) \quad (2)$$

其中， c_i 和 e_i 分别为影响个体 i 收入的环境因素和努力因素， μ 表示未观测到的随

机因素。由于努力因素难以量化，文献通常使用个体的受教育程度和就业选择作为努力水平的指代变量（罗良文、茹雪，2019）。所以，个体收入函数可表示为：

$$\ln(y_i) = \alpha C_i + \beta E_i + \mu_i \quad (3)$$

在上式中， C_i 表示可观测到的环境变量，包括个体特征（年龄、性别等）和家庭特征（家庭的社会经济地位、父母的受教育水平等）， E_i 为个体的努力变量， μ_i 为随机误差项， α 为环境因素对收入的直接影响系数。

在 Roemer（1998）对机会均等的定义中，如果环境与努力正交，则与环境相关的任何其他收入决定因素也应该理解成某种“环境”。基于上述假说，在 Trannoy et al.（2010）提出的模型基础上，我们引入个体的受教育程度和以个体入职的职业等为代表的社会经济地位（socioeconomic status，简称 SES）来构成一种“环境”，从事前角度来测算环境因素通过影响努力因素对收入的间接影响。环境对努力的影响可表示为：

$$E_i = \gamma C_i + \xi_i \quad (4)$$

式（4）中的 γ 表示个体的环境因素对努力变量的影响系数， ξ_i 为影响努力的残差部分，本文假设残差 ξ_i 和环境因素 C_i 相互独立。当式（3）中的努力变量 E 是个体的受教育程度， C_i 为环境因素时，式（3）变为：

$$\ln(y_i) = (\alpha + \beta\gamma)C_i + \delta_i \quad (5)$$

式（5）中， α 为环境因素对收入的直接影响系数， $\beta\gamma$ 为环境因素通过影响个体的受教育程度对收入的间接影响系数， $(\alpha + \beta\gamma)$ 为环境因素加上个体的受教育程度构成的新的“环境”对收入的影响系数。

当式（4）中的努力变量 E 是个体的社会经济地位时，式（4）可重新表示为 $E_i = \zeta C_i + \xi_i$ 。其中， C_i 为环境因素加上个体的受教育程度构成的新的“环境”，并假设残差 ξ_i 与环境因素 C_i 相互独立。将其带入式（3）可得：

$$\ln(y_i) = (\alpha + \beta\zeta)C_i + \varepsilon_i \quad (6)$$

式（6）中， $\beta\zeta$ 为环境因素通过影响个体的社会经济地位对收入的间接影响系数， $(\alpha + \beta\zeta)$ 为客观的外部环境对收入影响的总效应系数。

从式（4）到式（6）模型构建的依据在于个体都是先接受教育，再进入劳动力市场。所以，本文在式（3）的基础上先后引入个体的受教育水平和社会经济地位变量，分别根据这两种不同的效应来测算环境因素对机会不平等的间接效应。

（二）机会不平等度量指标的选择

Roemer & Trannoy（2016）指出，量化、排名和分解是应用机会均等分析的三种常见操作，但首先需要有一个合适的指标来衡量机会不平等的程度。Ferreira & Gignoux

(2011) 提出不平等指标应该满足对称性（匿名性）、Pigou-Dalton 转移性、尺度不变性、人口无关性。在诸多不平等度量指标中，只有广义熵指数满足以上四个性质并且还具可加可分解性，可将不平等分解成组间和组内两个部分。因此，本文使用平均对数偏差（Mean Log Deviation，简称 MLD）作为不平等指数，因为 MLD 属于广义熵指数，具有与路径无关的分解（Foster & Shneyerov, 2000）。使用基于群体人口份额的权重，对于收入分配 y （平均值 \bar{y} ）MLD 定义为：

$$I_{MLD}(y) = \frac{1}{n} \sum_{i=1}^n \ln\left(\frac{\bar{y}}{y_i}\right) \quad (7)$$

同时，MLD 可做如下分解：

$$I_{MLD}(y) = \sum_{t=1}^T p_t \ln(\bar{y}/\bar{y}_t) + \sum_{t=1}^T p_t I_{MLD}(y^t) \quad (8)$$

在式（8）中， p_t 为每组 t 内的人口份额， \bar{y}_t 为每组 t 内的收入均值。等式右边第一项表示组间差距，第二项表示为组内差距。有学者提出，组间部分的不平等也可以解释为平滑分布的不平等，其中，来自每组 t （在相同的环境下）的所有人都获得相同的收入 $y_i^t = y_i | c_i^t$ 。组内部分的不平等也可以解释为标准化的不平等，其中，各组之间的差异均已消除，只有组内的差异得以保留。

（三）机会不平等的 Shapley 分解

基于衡量机会不平等的目的，我们可将上文的式（3）、式（5）和式（6）式简化为：

$$\ln y_i = \psi C_i + \varepsilon_i \quad (9)$$

在式（9）中，系数 ψ 即包括环境因素对收入 y 的直接影响，也包括环境通过与环境相关的努力变量（如受教育程度和社会经济地位）对收入 y 的间接影响。通过对式（9）的最小二乘（OLS）回归分析，得到简化形式系数 ψ 的估计，就可以构建平滑分布的参数估计，如下所示：

$$\tilde{\mu}_i = \exp[\hat{\psi} C_i] \quad (10)$$

将 MLD 应用在平滑分布来计算 IO（组间的不平等），即 $IO = I_{MLD}(\tilde{\mu})$ 。

同样，组内的不平等等于均等化的环境 \bar{C} 加上未被观测到的个体变量，即：

$$\tilde{\phi}_i = \exp[\hat{\psi} \bar{C}_i + \varepsilon_i] \quad (11)$$

其中 $\hat{\varepsilon}_i = \ln y_i - \hat{\psi} C_i$ ，该分布中的不平等（残差不平等）可表示为： $RI = I_{MLD}(\tilde{\phi})$ ，即可观测到的环境因素无法解释的不平等。最终可将不平等分解为：

$$\underbrace{I_{MLD}(y)}_{I(y)} = \underbrace{I_{MLD}(\tilde{\mu})}_{IO} + \underbrace{I_{MLD}(\tilde{\phi})}_{RI} \quad (12)$$

在式（12）中，收入不平等 $I(y)$ 可以分解为组间的不平等 IO 和组内的不平等

RI, 根据 MLD 指数的分解可知, 由环境因素导致的机会不平等的绝对量为: $\theta^\alpha = I_{MLD}(\tilde{\mu})$, 而机会不平等相对量即在收入不平等中的占比可表示为: $\theta^r = I_{MLD}(\tilde{\mu}) / I_{MLD}(y)$ 。若式 (9) 为式 (3) 时, 则由环境集产生的机会不平等的绝对量为: $IO_C = I_{MLD}(\tilde{\mu})$; 若式 (9) 为加入教育变量的式 (5) 时, 由环境集产生的机会不平等的绝对量为: $IO_{EDU.C} = I_{MLD}(\tilde{\mu})$, 则个体受教育程度的间接效应绝对量为: $\theta_{EDU}^\alpha = IO_{EDU.C} - IO_C$, 相对量为: $\theta_{EDU}^r = \theta_{EDU}^\alpha / I(y)$; 若式 (9) 为加入社会经济地位变量的式 (6) 时, 由环境集产生的机会不平等的绝对量为: $IO_{SES.EDU.C} = I_{MLD}(\tilde{\mu})$, 则社会经济地位的间接效应绝对量为: $\theta_{SES}^\alpha = IO_{SES.EDU.C} - IO_{EDU.C}$, 相对量为: $\theta_{SES}^r = \theta_{SES}^\alpha / I(y)$; 环境集对机会不平等总的间接效应绝对量为: $\theta_{ind}^\alpha = IO_{SES.EDU.C} - IO_C$, 相对量为: $\theta_{ind}^r = \theta_{EDU}^r + \theta_{SES}^r$ 。

在根据上述分析计算得出机会不平等的值之后, 本文将基于 Shapley 分解法对机会不平等做进一步的分解, 以此来探讨在机会不平等的形成中, 可观测到的环境因素中哪些起着重要的作用; 同样, 我们也可探究当把与环境相关的努力变量编入环境因素中, 这些变量对机会不平等有多大的影响。基于对式 (3)、式 (5) 和式 (6) 的 OLS 回归分析, 我们首先估计环境变量所有可能排列的不等式的度量。然后, 我们计算每种环境变量对机会不平等的平均边际效应。

三 变量的选取与描述性统计

(一) 数据的来源与处理

本文所用的数据是由国家人口与计划生育委员会发布的 2010 年至 2017 年全国流动人口动态监测调查 (CMDS) 数据。CMDS 的调查对象是在流入地居住一个月以上, 非本区 (县、市) 户口的 15 周岁以上的流入人口。CMDS 以全国 31 个省 (区、市) 和新疆生产建设兵团流动人口年报数据为基本抽样框, 采取分层、多阶段、与规模成比例的方法进行抽样。同时, 样本量的分配不仅兼顾对全国、各省的代表性, 还增强了主要城市、均等化重点联系城市整体的代表性。所以, 该数据具有很强的可靠性和代表性。依据调查的技术文件, 我们以调查当年“五一”节前一周是否做过一个小时以上有收入的工作为识别变量, 剔除失业人员样本, 并选取流动原因是“务工”的样本, 就业身份限定为“雇员”。同时, 删除相关变量观察值缺失的样本。得到有效样本量分别为 2010 年 49784 个, 2012 年 66523 个, 2014 年 41755 个, 2017 年 60707 个^①。

^① 其中 2017 年数据来自当年 CMDS 第二轮调查数据 (基于流动人口问卷 A)。

（二）变量的选择与设计

收入（income）包括个人工资、奖金及其他福利性收入。流动人口中农村户籍个体占据大多数，其就业状态具有季节性和不稳定性，往往不能享受工资以外的收入。因此，本文以调查问卷中上个月的工资收入作为月收入变量。为保证数据的合理性并剔除极端值的影响，本文剔除了收入小于1%分位数和大于99%分位数的样本。

环境变量的选取：首先是个体特征变量。参考相关文献，本文选取年龄和性别变量。年龄 = 问卷调查年份 - 出生年份，同时加入年龄的二次项来考察年龄的变化趋势。性别变量设为虚拟变量，男性取1，女性取0。其次是家庭背景变量。由于数据中双亲变量混在一起，无法单独区分，且有效变量过少，本文选取户口性质和出生地变量。户口性质设为虚拟变量，非农业户口取1，农业户口取0；出生地变量根据国家统计局最新标准分为东中西部，以西部为基准，中部和东部分别设置虚拟变量。

努力变量的选取：本文选取个体受教育程度和以行业、职业等为代表的社会经济地位（SES）。教育程度设定为虚拟变量，未上学的取0，小学取1，初中取2，高中或中专取3，大专取5，本科取6，研究生以上取7。本文参考相关文献以个体就业的行业、职业、企业作为个人社会经济地位的代理变量。遵循岳希明等（2010）的划分方法，本文将20类行业分为垄断行业、竞争行业和其他三大类，并设置虚拟变量。其中垄断行业包括电煤水热生产供应、交通运输、仓储和邮政、信息传输、软件和信息技术服务、金融等，将其设置为参照组；竞争性行业包括建筑、批发零售、居民服务、修理和其他服务业等；其他行业包括农林牧渔业、采矿业、制造业、房地产、租赁和商务副业等。本文将企业按所有制性质分为国有企业、集体企业、私营企业和外资企业四大类，并设置虚拟变量。国有企业包括机关、事业单位、国有及国有控股企业等；集体企业包括集体独资企业、集体控股企业等；私营企业包括私营独资企业、私营控股企业等；外资企业包括外商合资企业、外资独资企业等，将其设置为参照组。本文将职业分为白领、蓝领和服务业三大类，并设置虚拟变量，将就职于国家机关、党群组织、企事业单位负责人、专业技术人员、公务员、办事人员和有关人员等岗位的划为白领；而蓝领包括农林牧渔、生产、运输、建筑等岗位；服务业包括商业、服务业人员等，将其设置为参照组。

（三）数据的描述性分析

表1的描述性统计显示，流动人口个体收入的均值不论是在全样本中还是在子样本中均呈现递增趋势。在全样本中，个体收入均值在2010年、2012年、2014年和2017年分别为1917.96元、2629.56元、3242.72元和4071.13元。收入中位数也呈现

相似的递增趋势。在户籍和性别的子样本中，城镇户籍流动人口的收入水平明显高于农村户籍流动人口，且男性的收入水平高于女性。农村户籍女性在收入方面处于户籍和性别双重弱势地位。在受教育程度上，从 2010 年到 2017 年，流动人口的平均受教育程度显著提高，尤其是只接受低层次教育（初中及其以下）者的比重从 2010 年的 67.12% 下降至 2017 年的 50.69%。同样，本文也发现流动人口中女性的平均受教育程度较低。在就业的选择上，以职业分类为例，服务业的从业比例有逐年上升的趋势，但增幅不大。有 5 成以上的女性流动人口从事服务业，男性流动人口则更多从事蓝领职业。在工资收入高、福利政策好、工作稳定有保障的白领岗位上，男性有明显的优势。

表 1 描述性统计

	2010 年			2012 年			2014 年			2017 年		
	全样本	男性	女性									
样本量	49784	28012	21772	66523	38431	28092	41755	25283	16472	60707	34641	26066
收入(元)												
均值	1917.96	2138.27	1634.51	2629.56	2892.20	2270.26	3242.72	3562.47	2751.93	4071.13	4543.89	3442.83
中位数	1800	2000	1500	2402	2700	2000	3000	3000	2500	3500	4000	3000
城镇	2202.46	2407.49	1944.27	3116.27	3344.35	2822.20	3849.85	4131.78	3423.59	4741.82	5227.50	4107.76
农村	1865.88	2089.53	1576.99	2522.98	2796.21	2144.29	3059.71	3392.19	2547.06	3854.17	4436.28	3267.34
最小值	600	600	600	800	800	800	1000	1000	1000	1000	1000	1000
标准差	852.22	892.58	702.09	1202.98	1244.51	1041.48	1579.71	1648.56	1325.37	2088.52	2188.48	1761.27
最大值	6000	6000	6000	9000	9000	9000	10000	10000	10000	15000	15000	15000
受教育程度												
均值	9.885	9.946	9.808	10.199	10.215	10.176	10.478	10.480	10.476	10.391	10.463	10.295
初中以下	12.13%	10.83%	13.80%	11.74%	10.38%	13.60%	7.31%	6.69%	8.27%	11.21%	9.46%	13.53%
初中	54.99%	56.11%	53.54%	49.95%	51.81%	47.39%	43.69%	45.43%	41.02%	39.48%	41.44%	36.87%
高中	23.26%	23.61%	22.81%	24.37%	24.88%	23.67%	40.06%	38.77%	42.03%	38.58%	38.79%	38.30%
高中以上	9.62%	9.44%	9.86%	13.95%	12.92%	15.32%	8.95%	9.12%	8.68%	10.74%	10.31%	11.30%
职业												
服务业	39.28%	30.65%	50.38%	39.46%	30.39%	51.86%	60.63%	51.05%	75.34%	43.32%	36.52%	52.35%
蓝领	42.02%	46.27%	36.56%	45.86%	53.13%	35.91%	23.74%	31.50%	11.83%	37.26%	43.04%	29.57%
白领	18.70%	23.09%	13.06%	14.69%	16.49%	12.23%	15.63%	17.45%	12.83%	19.43%	20.44%	18.07%

资料来源：根据 CMDS 数据计算得到。

表 2 报告了主要变量的相关性分析。结果表明，无论是 Pearson 检验还是 Spearman 检验，流动人口的个人收入与“环境”中的个体特征变量年龄（Age）、性别（Gender）、户籍（Household）和出生地（Region）显著相关，且与受教育程度（Edu）

和社会经济地位变量显著相关。另外，个人的受教育程度（Edu）不仅与“环境”变量均显著相关，还和社会经济地位变量显著相关。

表2 相关性分析

	Income	Age	Gender	Household	Region	Edu	Occupation	Industry	Enterprise
Income	1	0.071 ***	0.269 ***	0.126 ***	0.060 ***	0.173 ***	0.180 ***	-0.009 ***	0.031 ***
Age	0.038 ***	1	0.104 ***	0.042 ***	-0.019 ***	-0.261 ***	0.053 ***	0.001	0.098 ***
Gender	0.210 ***	0.109 ***	1	-0.001	-0.008 ***	0.021 ***	0.154 ***	-0.041 ***	0.074 ***
Household	0.146 ***	0.034 ***	-0.001	1	0.066 ***	0.381 ***	0.105 ***	-0.033 ***	0.125 ***
Region	0.068 ***	-0.024 ***	-0.009 ***	0.065 ***	1	0.142 ***	0.011 ***	-0.002	-0.033 ***
Edu	0.191 ***	-0.255 ***	0.022 ***	0.405 ***	0.143 ***	1	0.132 ***	-0.045 ***	0.084 ***
Occupation	0.127 ***	0.029 ***	0.144 ***	0.134 ***	0.020 ***	0.212 ***	1	0.424 ***	0.057 ***
Industry	-0.066 ***	-0.002	-0.056 ***	-0.046 ***	-0.008 ***	-0.051 ***	0.329 ***	1	-0.091 ***
Enterprise	0.029 ***	0.095 ***	0.073 ***	0.150 ***	-0.028 ***	0.138 ***	0.129 ***	-0.059 ***	1

注：Income 为个体的月收入水平；Age 为年龄；Gender 为性别，其中男性 = 1，女性 = 0；Household 代表户籍，城镇 = 1，农村 = 0；Region 代表出生地，西部 = 0，中部 = 1，东部 = 2；Edu 代表个体受教育程度；Occupation 代表职业，服务业 = 0，蓝领 = 1，白领 = 2；Industry 代表行业，垄断行业 = 0，竞争行业 = 1，其他行业 = 2；Enterprise 代表企业，外资企业 = 0，私人企业 = 1，集体企业 = 2，国有企业 = 3；表格上三角为 Pearson 相关系数，下三角为 Spearman 相关系数；*、**、*** 分别表示显著水平为 10%、5%、1%。

资料来源：根据 CMDS 数据计算得到。

四 实证结果与分析

（一）OLS 回归分析

本文基于模型（3）、（5）和（6），在可观测到的环境集的基础上，依次引入与环境因素相关的努力变量（教育变量和社会经济地位变量）进行 OLS 回归估计。结果如表 3 所示。以 2017 年为例，若把年龄看成经验的累积，其每增加一个单位就会带来 5% 左右的收入增长且系数在 1% 水平上显著，而年龄的平方与收入呈显著的负相关，表明年龄与收入是“倒 U 型”关系。性别间收入差距明显高于户籍间收入差距，正与表 1 的描述性统计相符合。在回归结果中，男性的收入水平比女性高出 20% 以上，城镇户籍流动人口的收入水平比农村户籍流动人口高出 5% 以上，差异均在 1% 水平上显著。

在出生地上，东部地区和中部地区相对于西部地区有明显的优势，并且东部出生个体的收入高于中部出生个体，其估计系数均在 1% 水平上显著。以 2017 年为例，出生于东部的流动人口收入水平比出生于西部的流动人口高 6% 左右，而出生于中部的流动人口收入水平比出生于西部的流动人口高 5% 左右。在个人受教育程度方面，每增加

表 3 OLS 回归分析

年份	基于式(3)			基于式(5)			基于式(6)			
	2010	2012	2014	2010	2012	2014	2010	2012	2014	2017
年龄	0.050 ^{***} (38.25)	0.054 ^{***} (47.40)	0.068 ^{***} (45.25)	0.048 ^{***} (37.46)	0.051 ^{***} (45.82)	0.062 ^{***} (41.74)	0.054 ^{***} (47.82)	0.044 ^{***} (40.91)	0.053 ^{***} (36.61)	0.049 ^{***} (44.44)
年龄的平方	-0.001 ^{***} (-37.16)	-0.001 ^{***} (-46.12)	-0.001 ^{***} (-45.50)	-0.001 ^{***} (-34.41)	-0.001 ^{***} (-41.92)	-0.001 ^{***} (-39.72)	-0.001 ^{***} (-48.59)	-0.001 ^{***} (-37.76)	-0.001 ^{***} (-35.21)	-0.001 ^{***} (-45.19)
性别	0.261 ^{***} (74.75)	0.252 ^{***} (81.54)	0.255 ^{***} (63.99)	0.252 ^{***} (73.16)	0.245 ^{***} (80.69)	0.252 ^{***} (64.50)	0.268 ^{***} (79.12)	0.202 ^{***} (66.53)	0.204 ^{***} (52.21)	0.246 ^{***} (72.67)
户籍	0.130 ^{***} (27.16)	0.161 ^{***} (40.62)	0.178 ^{***} (38.24)	0.034 ^{***} (6.48)	0.051 ^{***} (11.37)	0.092 ^{***} (18.50)	0.067 ^{***} (16.11)	0.052 ^{***} (12.00)	0.080 ^{***} (16.57)	0.068 ^{***} (16.46)
中部	0.066 ^{***} (15.04)	0.052 ^{***} (13.78)	0.022 ^{***} (4.50)	0.059 ^{***} (13.57)	0.039 ^{***} (10.50)	0.007 ^{***} (1.54)	0.049 ^{***} (11.74)	0.037 ^{***} (10.44)	0.004 ^{***} (0.86)	0.048 ^{***} (11.76)
东部	0.085 ^{***} (18.24)	0.082 ^{***} (19.98)	0.073 ^{***} (13.85)	0.063 ^{***} (13.59)	0.056 ^{***} (13.95)	0.045 ^{***} (8.69)	0.065 ^{***} (14.39)	0.052 ^{***} (13.41)	0.038 ^{***} (7.51)	0.063 ^{***} (14.10)
教育	-	-	-	0.089 ^{***} (40.78)	0.091 ^{***} (51.98)	0.121 ^{***} (44.09)	0.122 ^{***} (55.94)	0.081 ^{***} (45.27)	0.097 ^{***} (34.89)	0.113 ^{***} (48.88)
竞争行业	-	-	-	-	-	-	-	-0.083 ^{***} (-11.84)	-0.104 ^{***} (-16.64)	-0.074 ^{***} (-10.85)
其他	-	-	-	-	-	-	-	-0.100 ^{***} (-17.62)	-0.106 ^{***} (-16.58)	-0.083 ^{***} (-12.80)

续表

年份	2010	2012	2014	2017	2010	2012	2014	2017	2010	2012	2014	2017
	基于式(3)				基于式(5)				基于式(6)			
私人企业	-	-	-	-	-	-	-	-	-0.086 ^{***} (-14.59)	-0.087 ^{***} (-17.36)	-0.113 ^{***} (-13.15)	-0.042 ^{***} (-6.70)
集体企业	-	-	-	-	-	-	-	-	-0.080 ^{***} (-7.96)	-0.075 ^{***} (-8.46)	-0.123 ^{***} (-9.24)	-0.005 ^{***} (-0.63)
国有企业	-	-	-	-	-	-	-	-	-0.075 ^{***} (-9.46)	-0.134 ^{***} (-21.12)	-0.182 ^{***} (-18.66)	-0.118 ^{***} (-15.56)
蓝领	-	-	-	-	-	-	-	-	0.144 ^{***} (28.56)	0.195 ^{***} (49.36)	0.168 ^{***} (33.04)	0.124 ^{***} (28.28)
白领	-	-	-	-	-	-	-	-	0.254 ^{***} (46.98)	0.256 ^{***} (48.88)	0.200 ^{***} (31.79)	0.178 ^{***} (34.04)
常数项	6.448 ^{***} (305.63)	6.678 ^{***} (353.90)	6.649 ^{***} (263.58)	7.081 ^{***} (344.47)	6.173 ^{***} (282.92)	6.402 ^{***} (332.75)	6.319 ^{***} (245.23)	6.648 ^{***} (309.36)	6.422 ^{***} (275.64)	6.625 ^{***} (322.55)	6.735 ^{***} (243.23)	6.805 ^{***} (289.99)
观测值个数	49784	66523	41755	60707	49784	66523	41755	60707	49784	66523	41755	60707
调整的R2	0.147	0.143	0.168	0.169	0.175	0.177	0.205	0.210	0.244	0.239	0.261	0.238

注：*、**、*** 分别表示显著水平为10%、5%、1%；括号内为t统计量。
资料来源：根据 CMDS 数据计算得到。

一个单位的受教育程度使得流动人口有至少 7% 的收入增长。社会经济地位变量的估计系数均在 1% 水平上显著，其中白领的收入水平比服务业从业者高出至少 17.8%，蓝领的收入水平比服务业从业者高出至少 12.4%。根据表 1 可知，女性流动人口有 5 成以上从事服务业，这也从职业就业方面解释了流动人口性别收入差距之大的原因所在。

(二) 机会不平等的测算与分解分析

以基于式 (3) 的回归估计系数与式 (10) 构建的反事实收入模型为基础，本文分年度测算了全样本、分户籍、分性别、分出生地和出生序列的收入不平等指数、机会不平等的绝对值和机会不平等的相对值，具体结果如表 4 和表 5 所示。

表 4 收入不平等和机会不平等的绝对值

年份	收入不平等(MLD 指数)				机会不平等(MLD 指数)			
	2010	2012	2014	2017	2010	2012	2014	2017
全样本	0.08802	0.09121	0.09904	0.10856	0.012523	0.012708	0.015737	0.017322
男性	0.08078	0.08338	0.09272	0.09962	0.006767	0.007062	0.009541	0.009291
女性	0.07740	0.08523	0.08898	0.09882	0.003466	0.006273	0.009287	0.007632
城镇	0.10541	0.11586	0.11529	0.12874	0.010352	0.009800	0.012715	0.013468
农村	0.08262	0.08162	0.08778	0.09659	0.011399	0.010762	0.012333	0.014843
东部	0.08885	0.09294	0.10345	0.11485	0.012442	0.012338	0.016637	0.016215
中部	0.08679	0.09057	0.09822	0.10766	0.012424	0.012604	0.016204	0.017563
西部	0.08656	0.08787	0.08810	0.09583	0.010972	0.012046	0.011779	0.013746
50 后	0.09821	0.10428	0.10460	0.12092	0.009758	0.009212	0.006812	0.013917
60 后	0.08897	0.09457	0.10378	0.10182	0.016614	0.016233	0.017334	0.013477
70 后	0.09383	0.09776	0.10607	0.11063	0.016196	0.015571	0.020619	0.018615
80 后	0.08210	0.08642	0.09700	0.10584	0.006594	0.007761	0.009913	0.012488

资料来源：根据 CMDS 数据计算得到。

在全样本中，收入不平等指数和机会不平等指数的变化趋势大体一致，二者均有所增长，但增幅并不大。在表 5 中，在缺乏部分环境变量（父母的受教育程度、职业和单位性质、家庭的社会经济地位等）和部分环境变量不可观测的情况下，本文基于现有的环境集对 2010 年到 2017 年做机会不平等下限估计，得出机会不平等在收入不平等中的占比分别为 14.23%、13.93%、15.89% 和 15.96%。马占利和邹薇（2018）使用中国家庭收入调查（CHIP）数据，通过参数法测算的机会不平等的相对值在 13.04% 至 19.16% 之间。

从性别方面来看，男性和女性的收入不平等指数比较接近，而在机会不平等指数和机会不平等在收入不平等中的占比上，男性略微高于女性。分城乡看，农村户籍流

动人口的收入不平等指数低于城镇户籍流动人口，2010 年城镇户籍样本收入不平等指数为 0.10541，农村户籍样本则为 0.08262，2017 年城镇户籍样本收入不平等指数为 0.12874，略有升高，同时期的农村户籍样本收入不平等指数为 0.09659。

分出生地看，东部地区出生样本的收入不平等指数和机会不平等指数最高。以 2017 的数据为例，东部组别的收入不平等指数为 0.11485，高于中部组别的 0.10766 和西部组别的 0.09583。三个地区的机会不平等指数均呈现递增趋势且中部地区增幅最大。在机会不平等的相对值中，东部、中部和西部三个组别的机会不平等在收入不平等的占比大致在 13% 至 16% 之间，基本与全样本的比例范围保持一致。

从出生队列来看，不同出生年代的差异性比较明显。80 后的收入不平等指数和机会不平等指数最低，从 2010 年到 2017 年，机会不平等的占比从 8.03% 逐步增加到 11.80%，但相比于其他出生队列依然很低。对此，史新杰等（2018）指出，80 后正好出生于改革开放后，相对于祖辈和父辈，他们从出生就生活在更加开放和公平的社会环境中。60 后和 70 后的机会不平等指数最高，机会不平等在收入不平等中的占比最高达到 19% 左右。李莹等（2019）指出，60 后和 70 后在改革开放初期进入劳动力市场，正是新的利益格局开始形成、阶层开始分化的时期，在这个时期整个社会的机会不平等程度正在显著提高。

表 5 机会不平等的相对值

单位：%

年份	2010	2012	2014	2017
全样本	14.23	13.93	15.89	15.96
男性	8.38	8.47	10.29	9.33
女性	4.48	7.36	10.44	7.72
城镇	9.82	8.46	11.03	10.46
农村	13.80	13.19	14.05	15.37
东部	14.00	13.28	16.08	14.12
中部	14.32	13.92	16.50	16.31
西部	12.68	13.71	13.37	14.34
50 后	9.94	8.83	6.51	11.51
60 后	18.67	17.17	16.70	13.24
70 后	17.26	15.93	19.44	16.83
80 后	8.03	8.98	10.22	11.80

资料来源：根据 CMDS 数据计算得到。

接着，我们基于 Shapley 分解方法分解各种环境因素对收入机会不平等的贡献。在模型（3）回归基础上，我们计算了 2010 年至 2017 年环境变量对机会不平等的贡献

度。为了避免扭曲收入变量的含义与分布，此处被解释变量使用的是收入本身，而非收入的对数。计算结果如表 6 至表 9 所示。

在全样本中，个体特征的环境因素在很大程度上解释了机会不平等，而出生地对机会不平等的贡献很小。在个体特征的环境因素中，性别因素、年龄因素和户籍因素对机会不平等的贡献度依次减弱。其中，性别是机会不平等的首要因素，其贡献率在 2010 年为 66.39%，在 2017 年为 54.15%，呈下降趋势。

在以环境因素分组的子样本中，各个群体在各年的数据中，每个环境因素的贡献度分布均大体一致。然而，各因素的贡献率在不同性别之间有明显的差别。其中，男性群体中年龄因素对机会不平等的贡献率在 66.47% ~ 77.56% 之间，户籍因素次之。从表 1 可以看出，男性流动人员有四成左右就业于以建筑业、采掘业为代表的蓝领职业，这些重体力劳动与年龄有着很强的关联。对于女性群体，机会不平等贡献因素中户籍占据着主要地位，但呈现逐步下降的趋势，从 2012 年的 57.57% 下降到 2017 年的 41.92%。

从城乡分组看，性别因素造成的机会不平等均占很大比重。其中，农村户籍群组性别因素的贡献率远高于城镇户籍群组。通过上述分析，我们可以观察到，按性别分组，女性群组中户籍因素是机会不平等生成的主要因素，按户籍分组，农村户籍群组中性别因素又是机会不平等生成的主要因素，这与表 1 中的描述相一致，即女性流动人口在劳动力市场上处于来自户籍和性别的双重弱势地位。在出生地分组和出生队列分组中，性别因素依然是产生机会不平等的主要因素，这一情况在 60 后和 70 后中表现得更为突出。

表 6 机会不平等的分解——基于 Shapley 值分解 (CMDS-2010)

	IO	年龄(占比%)		性别(占比%)		户籍(占比%)		区域(占比%)	
全样本	0.012523	0.002420	19.33	0.008313	66.39	0.001268	10.13	0.000521	4.16
男性	0.006767	0.005248	77.56	-	-	0.000894	13.21	0.000625	9.23
女性	0.003466	0.001105	31.89	-	-	0.001913	55.19	0.000448	12.92
城镇	0.010352	0.003311	31.98	0.005963	57.60	-	-	0.001078	10.41
农村	0.011399	0.002220	19.47	0.008794	77.14	-	-	0.000385	3.38
东部	0.012442	0.002095	16.84	0.008139	65.42	0.002208	17.75	-	-
中部	0.012424	0.002533	20.39	0.009070	73.00	0.000821	6.61	-	-
西部	0.010972	0.002890	26.34	0.007278	66.33	0.000804	7.32	-	-
50 后	0.009758	-	-	0.006867	70.38	0.002088	21.40	0.000802	8.22
60 后	0.016614	-	-	0.015002	90.30	0.001331	8.01	0.000281	1.69
70 后	0.016196	-	-	0.014482	89.42	0.001361	8.41	0.000353	2.18
80 后	0.006594	-	-	0.004454	67.54	0.001358	20.60	0.000782	11.86

资料来源：根据 CMDS2010 数据计算得到。

表 7 机会不平等的分解——基于 Shapley 值分解 (CMDS - 2012)

	IO	年龄(占比%)		性别(占比%)		户籍(占比%)		区域(占比%)	
全样本	0.012708	0.002642	20.79	0.007359	57.91	0.002188	17.22	0.000518	4.08
男性	0.007062	0.005305	75.13	-	-	0.001413	20.01	0.000343	4.86
女性	0.006273	0.001771	28.23	-	-	0.003612	57.57	0.000890	14.19
城镇	0.009800	0.003252	33.19	0.004326	44.14	-	-	0.002222	22.68
农村	0.010762	0.002343	21.77	0.008175	75.96	-	-	0.000244	2.27
东部	0.012338	0.002263	18.34	0.005895	47.78	0.004181	33.88	-	-
中部	0.012604	0.002798	22.20	0.007769	61.64	0.002037	16.16	-	-
西部	0.012046	0.002919	24.23	0.008533	70.84	0.000594	4.93	-	-
50后	0.009212	-	-	0.003972	43.12	0.003596	39.04	0.001644	17.84
60后	0.016233	-	-	0.014507	89.37	0.001647	10.15	0.000079	0.49
70后	0.015571	-	-	0.013261	85.17	0.001955	12.55	0.000355	2.28
80后	0.007761	-	-	0.004259	54.88	0.002701	34.81	0.000801	10.32

资料来源：根据 CMDS2012 数据计算得到。

表 8 机会不平等的分解——基于 Shapley 值分解 (CMDS - 2014)

	IO	年龄(占比%)		性别(占比%)		户籍(占比%)		区域(占比%)	
全样本	0.015737	0.004259	27.06	0.007594	48.26	0.003319	21.09	0.000566	3.59
男性	0.009541	0.006786	71.13	-	-	0.002203	23.09	0.000552	5.79
女性	0.009287	0.003143	33.84	-	-	0.005521	59.45	0.000623	6.71
城镇	0.012715	0.006195	48.72	0.004446	34.97	-	-	0.002074	16.31
农村	0.012333	0.003374	27.36	0.008730	70.79	-	-	0.000228	1.85
东部	0.016637	0.003668	22.05	0.007138	42.90	0.005832	35.05	-	-
中部	0.016204	0.005001	30.86	0.008171	50.42	0.003032	18.71	-	-
西部	0.011779	0.003409	28.94	0.007194	61.07	0.001176	9.99	-	-
50后	0.006812	-	-	0.005605	82.28	0.000759	11.14	0.000448	6.58
60后	0.017334	-	-	0.016275	93.89	0.001025	5.91	0.000034	0.20
70后	0.020619	-	-	0.015748	76.38	0.004518	21.91	0.000353	1.71
80后	0.009913	-	-	0.005068	51.13	0.003938	39.73	0.000907	9.14

资料来源：根据 CMDS2014 数据计算得到。

表 9 机会不平等的分解——基于 Shapley 值分解 (CMDS - 2017)

	IO	年龄(占比%)		性别(占比%)		户籍(占比%)		区域(占比%)	
全样本	0.017322	0.004526	26.13	0.009380	54.15	0.002516	14.53	0.000899	5.19
男性	0.009291	0.006175	66.47	-	-	0.002096	22.56	0.001020	10.97

续表

	IO	年龄(占比%)		性别(占比%)		户籍(占比%)		区域(占比%)	
女性	0.007632	0.003636	47.64	-	-	0.003199	41.92	0.000796	10.44
城镇	0.013468	0.004835	35.90	0.007828	58.12	-	-	0.000805	5.98
农村	0.014843	0.004244	28.60	0.009932	66.91	-	-	0.000666	4.49
东部	0.016215	0.003698	22.81	0.009264	57.13	0.003253	20.06	-	-
中部	0.017563	0.005457	31.07	0.010190	58.02	0.001916	10.91	-	-
西部	0.013746	0.003616	26.30	0.008258	60.08	0.001872	13.62	-	-
50 后	0.013917	-	-	0.005945	42.72	0.006619	47.56	0.001353	9.72
60 后	0.013477	-	-	0.011019	81.76	0.001968	14.60	0.000491	3.64
70 后	0.018615	-	-	0.016022	86.07	0.002015	10.83	0.000578	3.11
80 后	0.012488	-	-	0.008610	68.95	0.002810	22.50	0.001068	8.55

资料来源：根据 CMDS2017 数据计算得到。

(三) 环境对机会不平等的间接效应

正如上文所述, Roemer (1998) 提出, 如果努力变量和环境正交, 则可把受环境影响的努力变量看成新的环境变量, 构建出一个新的环境集。考虑到环境因素与努力因素的相关性 (如表 2 所示), 本文在模型 (3) 的基础上分别引入与环境因素相关的教育变量和社会经济地位变量, 并分别计算得出环境通过影响两个变量对机会不平等的间接效应的绝对量和相对量。

如表 10 和表 11 所示, 基于式 (5) 和式 (6) 测算的机会不平等 MLD 指数在全样本中均有明显的上升趋势。引入教育变量后, 机会不平等的绝对值从 2010 年的 0.014993 增加到 2017 年的 0.021577, 再引入社会经济地位变量后, 机会不平等的绝对值从 2010 年的 0.020703 增加到 2017 年的 0.024297。若分年份来看, 在 2010 年, 在逐次引入教育变量和社会经济地位变量后, 机会不平等的绝对量从 0.012523 增长到 0.014993 和 0.020703; 同样, 在 2017 年, 机会不平等的绝对量从 0.017322 增长到 0.021577 和 0.024297。这表明, 环境集通过教育和社会经济地位间接增加了机会不平等。

通过表 11 则可以看出, 在全样本中, 环境集通过教育对机会不平等的间接影响逐渐增强, 而环境集通过社会经济地位对机会不平等的间接影响逐渐减弱。在 2010 年, 教育对机会不平等的作用仅有 2.8%, 低于社会经济地位的 6.49%, 而在 2017 年, 社会经济地位对机会不平等的贡献下降至 2.5%, 教育则增长至 3.92%。

在异质性分组中, 女性流动人口面临的来自教育的收入不公大于男性, 即使两者

在教育对机会不平等间接效应的相对值上差别不大。然而，在环境对机会不平等的间接效应中，相较于教育间接效应，女性由入职的行业、职业等为代表的社会经济地位所带来的收入不公更为明显。在2010年，社会经济地位间接效应的相对值于女性为8.19%，于男性则为6.23%。农村户籍流动人口社会经济地位间接效应的相对值在2010年和2012年高于城镇户籍流动人口，但在2014年和2017年低于城镇户籍流动人口。

从地区分组来看，东部组别教育的间接效应大于中部和西部组别，而在社会经济地位的间接效应上三者差别不大，且变化趋势一致。在不同出生队列分组中，50后社会经济地位间接效应的相对值明显高于其他组别，且始终保持着较高值，在2010年50后社会经济地位间接效应的相对值为10.19%，增长至2012年的18.93%，在2017年依旧有13.89%的比例。同样，60后社会经济地位间接效应的相对值虽低于50后，却也维持在较高水平，从2010年的7.28%微增至2017年的7.76%，区别于70后和80后社会经济地位间接效应递减的趋势。这表明50后和60后流动人口群体在劳动力市场始终面对着来自职业、行业等带来的收入不公平，近年来这种情况并没有得到明显的改善。根据表10和表11，在考虑到环境通过影响努力因素对机会不平等的间接效应后，流动人口机会不平等的相对值在22.38%至24.61%之间。

表10 机会不平等的绝对值 (MLD 指数)

年份	2010		2012		2014		2017	
	(5)	(6)	(5)	(6)	(5)	(6)	(5)	(6)
全样本	0.014993	0.020703	0.015791	0.020965	0.019313	0.024370	0.021577	0.024297
男性	0.008730	0.013766	0.009041	0.013808	0.012332	0.017644	0.013201	0.015796
女性	0.006018	0.012357	0.010223	0.015872	0.013381	0.017246	0.012014	0.014828
城镇	0.022767	0.027555	0.022982	0.028693	0.024466	0.030328	0.027440	0.030827
农村	0.012482	0.018002	0.012106	0.016744	0.014093	0.018340	0.016995	0.019207
东部	0.016064	0.020774	0.017731	0.022045	0.023037	0.027763	0.024755	0.027090
中部	0.014621	0.020987	0.015761	0.021512	0.019812	0.025090	0.021178	0.023816
西部	0.012692	0.018620	0.013243	0.017996	0.013181	0.017826	0.015631	0.018820
50后	0.011317	0.021319	0.011787	0.031526	0.008548	0.027765	0.018412	0.035208
60后	0.018036	0.024512	0.017723	0.024670	0.019266	0.029278	0.015172	0.023070
70后	0.019172	0.024272	0.018927	0.024508	0.025092	0.031040	0.022946	0.028303
80后	0.008892	0.015337	0.010821	0.015923	0.013541	0.018458	0.016537	0.018507

资料来源：根据 CMDS 数据计算得到。

表 11 环境集对机会不平等生成的间接效应 (相对值)

单位: %

	教育的间接效应				社会经济地位 (SES) 的间接效应			
	2010	2012	2014	2017	2010	2012	2014	2017
全样本	2.80	3.38	3.61	3.92	6.49	5.68	5.11	2.50
男性	2.43	2.37	3.01	3.92	6.23	5.72	5.73	2.61
女性	3.30	4.63	4.60	4.44	8.19	6.63	4.34	2.85
城镇	11.78	11.38	10.19	10.85	4.54	4.93	5.09	2.64
农村	1.31	1.64	2.00	2.22	6.68	5.68	4.84	2.30
东部	4.08	5.80	6.19	7.43	5.30	4.64	4.57	2.04
中部	2.62	3.48	3.67	3.36	7.33	6.35	5.37	2.45
西部	1.98	1.36	1.59	1.97	6.85	5.41	5.27	3.33
50 后	1.58	2.47	1.66	3.72	10.19	18.93	18.37	13.89
60 后	1.60	1.57	1.86	1.66	7.28	7.35	9.65	7.76
70 后	3.17	3.43	4.22	3.91	5.44	5.71	5.60	4.84
80 后	2.80	3.54	3.74	3.82	7.85	5.91	5.07	1.87

资料来源: 根据 CMDS 数据计算得到。

接下来, 本文基于 Shapley 分解法计算在引入教育变量和社会经济地位变量后, 各环境因素对收入机会不平等的贡献。为了避免扭曲收入变量的含义与分布, 此处被解释变量使用的是收入本身, 而非收入的对数。最终的结果如表 12 所示。

首先, 在引入教育变量后, 性别因素依然是造成机会不平等的最主要的因素。在 2010 年, 性别因素对机会不平等的贡献度为 53.79%, 而户籍因素仅为 4.71%, 说明在流动人口中引起收入不平等的因素中, 性别而非户籍是主要原因。从纵向来看, 教育因素对机会不平等的贡献度逐年增加, 从 2010 年的 21.67% 逐步增加到 2017 年的 30.44%。这也与上述分析中教育变量间接效应的相对值在逐年增加的趋势一致。

进一步引入社会经济地位变量后, 各因素对机会不平等的贡献出现很大的变化, 社会经济地位中的职业变量对机会不平等的贡献从 2010 年的 37.43% 递减至 2014 年的 32.02%, 却依然是各年度中对机会不平等贡献度最大的因素。在 2017 年, 社会经济地位中的职业变量对机会不平等的贡献度下降至 18.26%, 这一下降幅度也与上述分析中社会经济地位间接效应的相对值逐年较大幅下降是一致的。分解也表明, 环境集通过影响社会经济地位变量对机会不平等的间接效应更多来自于职业所带来的收入不平等。教育对机会不平等的贡献从 2010 年的 13.49% 逐年增加至 2017 年的 23.73%, 与上述实证结果一致。在新构建的“环境集”中, 年龄、性别、户籍和出生地对机会不平等的总贡献度也从 2010 年的 47.24% 增加至 2017 年的 56.89%。

表 12 机会不平等指数分解——基于 Sharply 值分解 (CMDS 2010–2017)

	IO	个体特征		家庭背景		教育	社会经济地位 (SES)		
		年龄	性别	户籍	出生地	教育	企业	行业	职业
Panel A: 基于式(5)									
2010	0.014993	17.28%	53.79%	4.71%	2.56%	21.67%	-	-	-
2012	0.015791	17.81%	45.53%	7.88%	2.35%	26.44%	-	-	-
2014	0.019313	20.43%	38.92%	11.11%	2.09%	27.45%	-	-	-
2017	0.021577	18.19%	41.21%	7.16%	3.00%	30.44%	-	-	-
Panel B: 基于式(6)									
2010	0.020703	10.32%	32.39%	3.01%	1.52%	13.49%	0.21%	1.61%	37.43%
2012	0.020965	11.44%	28.40%	5.20%	1.69%	16.94%	0.75%	1.23%	34.35%
2014	0.024370	13.84%	25.18%	7.44%	1.51%	17.63%	0.97%	1.41%	32.02%
2017	0.024297	14.83%	33.67%	5.75%	2.64%	23.73%	0.52%	0.62%	18.26%

资料来源：根据 CMDS 数据计算得到。

五 环境对机会不平等间接效应的稳健性检验

在上一部分，我们分别引入个体受教育程度变量和社会经济地位变量到环境集的回归方程中，以此来测算环境因素通过影响努力变量对机会不平等的间接效应。这种方法的测度结果是否准确？比如在测度教育的间接影响时，城市的相对值是高于农村的，这一结果有悖于有关学者的研究结论（李莹等，2019）。同时，上述测算也忽略了因素间的因果关系，社会经济地位的职业类别、入职的行业等不仅受环境集的影响，而且还受个体受教育程度的影响。那么，三者之间的因果关系会不会对上述结果产生影响呢？对此，我们借鉴 Palomino et al. (2019) 提出的多层次反事实模型，进一步探究机会不平等间接效应的生成，并以此对上一部分的实证结果做稳健性检验。

Palomino et al. (2019) 认为环境不会直接作用于未来的个体收入，而是通过某些中间变量（如教育和职业类别等）影响个人的收入。通过上文对机会不平等 IO 的测算可知，平滑分布 $\tilde{\mu}_i$ 表示个体的总收入中由可观测的环境集 C_i 所解释的部分。如果令 $y_i^c = \mu_i$ ，则 y_i^c 为机会不平等测度中可观测的环境集 C_i 的反事实收入。在讨论机会不平等的间接效应作用机制之前，我们假定 y_i^c 是一组可观测到的且受环境集 C_i 影响的中间变量 Z_i 的函数，其函数形式可表示为：

$$y_i^c = f(Z_i, v_i) \quad (13)$$

其中 v_i 为观测到的中间变量和个体收入的随机组成部分。在第二部分的模型构建中已经指出 y_i^c 不仅包含可观测环境集的直接影响，还包含环境集通过个人受教育程度变量和社会经济地位变量的间接影响。同时，由于个体在生命周期内先接受教育然后进入市场，我们首先考虑个人的受教育程度作为第一效应，则有：

$$\ln y_i^c = \eta \text{EDU}_i + v_i \quad (14)$$

其中， EDU_i 为个人的受教育程度。通过对上式的 OLS 回归分析，可以得到教育系数的估计。同时，这个回归系数也可以用来获得一个“过度平滑”的收入分配，在这个分配中，所有具有相同教育程度的个体都将获得相同环境集下的条件收入，则可以根据平滑分布构建反事实收入为：

$$y_i^{c, \text{EDU}} = \exp[\hat{\eta} \text{EDU}_i] \quad (15)$$

其中， $y_i^{c, \text{EDU}}$ 为与个人受教育水平相关的环境条件下的收入的部分，即环境通过个人受教育程度间接影响收入的部分，所以，个体教育的机会不平等效应用 MLD 可以表示为： $IO_{\text{EDU}} = I_{\text{MLD}}(y_i^{c, \text{EDU}})$ 。类似地， v_i 的 MLD 可解释为不能被个人教育解释的机会不平等部分，即： $I_{\text{MLD}}(y_i^{c, \overline{\text{EDU}}}) = I_{\text{MLD}}(\exp[\hat{v}_i]) = I_{\text{MLD}}(\hat{\eta} \overline{\text{EDU}}_i + \hat{v}_i)$ 。

因此，机会不平等有如下的分解：

$$\frac{I_{\text{MLD}}(y_i^c)}{IO} = \frac{I_{\text{MLD}}(y_i^{c, \text{EDU}})}{IO_{\text{EDU}}} + \frac{I_{\text{MLD}}(y_i^{c, \overline{\text{EDU}}})}{IO_{\overline{\text{EDU}}}} \quad (16)$$

式 (16) 右边第一部分表示不同教育程度组间的不平等，即个人教育的间接效应；右边第二部分表示相同教育程度下组内的不平等，既包含环境集对机会不平等的直接效应，也包含环境集通过个人社会经济地位等因素对机会不平等的间接效应。

接着，我们进一步测算个体社会经济地位的间接效应。根据上述分析， $y_i^{c, \overline{\text{EDU}}}$ 可进一步分解成由个体的社会经济地位解释的部分和由其他因素引起的部分，则有：

$$\ln y_i^{c, \overline{\text{EDU}}} = \gamma \text{SES}_i + \xi_i \quad (17)$$

其中， SES_i 为个体职业类别等为代表的社会经济地位变量， ξ_i 为残差项。通过对式 (17) 的 OLS 回归估计，则可得到如下的反事实收入：

$$y_i^{c, \text{SES/EDU}} = \exp[\hat{\gamma} \text{SES}_i] \quad (18)$$

$$y_i^{c, \text{OTh}} = \exp[\hat{\xi}_i] = \exp[\hat{\gamma} \overline{\text{SES}}_i + \hat{\xi}_i] \quad (19)$$

因此，我们将 (16) 式机会不平等 IO 中不受个人教育作用的部分 $IO_{\overline{\text{EDU}}}$ 分解成个体社会经济地位效应 IO_{SES} 和其他因素的效应 IO_{OTh} 。

综上，我们对式 (16) 做进一步的分解，得到机会不平等 IO 的最终分解式：

$$\frac{I_{\text{MLD}}(y_i^c)}{IO} = \frac{I_{\text{MLD}}(y_i^{c, \text{EDU}})}{IO_{\text{EDU}}} + \frac{I_{\text{MLD}}(y_i^{c, \text{SES/EDU}})}{IO_{\text{SES}}} + \frac{I_{\text{MLD}}(y_i^{c, \text{OTh}})}{IO_{\text{OTh}}} \quad (20)$$

则个人教育的间接效应的相对值为： $IO_{EDU}^r = IO_{EDU}/IO$ ，个人社会经济地位的间接效应的相对值为： $IO_{SES}^r = IO_{SES}/IO$ 。

如表 13 和表 14 所示，在全样本中，环境通过影响教育对机会不平等的间接效应的绝对值从 2010 年的 0.00041 逐年增长至 2017 年的 0.00146，其相对值也从 2010 年的 3.27% 逐年增长至 2017 年的 8.43%，变化趋势与表 11 的结果一致。同样，社会经济地位间接效应的绝对值从 2010 年的 0.00095 下降至 2017 年的 0.00035，其相对值也从 2010 年的 7.59% 下降至 2017 年的 2.02%，与前述结果基本一致。

在异质性分组中，女性的教育间接效应大于男性。以 2017 年为例，男性教育间接效应的绝对值为 0.00109，女性为 0.00197；男性教育间接效应的相对值为 11.73%，女性为 25.81%。这与有关学者得出的结论相似（李莹、吕光明，2019）。城镇户籍和农村户籍组别在教育间接效应上的差别不大。以 2017 年为例，城镇户籍组的教育间接效应相对值为 1.78%，农村户籍组则为 3.71%，这一结果不同于表 11。

在出生地分组中，中部和西部地区教育间接效应的绝对值和相对值均呈现递增的趋势，而东部地区则呈现先增后减的趋势。在 2017 年，中部地区教育间接效应的相对值为 7.8%，大于西部地区和东部地区。在出生队列的分组中，50 后、60 后和 70 后的教育间接效应相对值较高，80 后则是最低的。以 2017 年为例，50 后的教育间接效应相对值为 17.1%，60 后为 13.65%，均大于 70 后和 80 后，50 后教育间接效应相对值比 80 后高出 11 个百分点。这表明在劳动力市场上，50 后和 60 后流动人口受到教育机会不平等的影响较大，并且这种不公平的影响并没有减弱，反而在增强。在社会经济地位的间接效应上，各个分组的相对值均保持与表 11 一致的分布和变化趋势。

表 13 环境集对机会不平等的间接效应（绝对值）

年份	教育的间接效应				社会经济地位 (SES) 的间接效应			
	2010	2012	2014	2017	2010	2012	2014	2017
全样本	0.00041	0.00067	0.00111	0.00146	0.00095	0.00074	0.00008	0.00035
男性	0.00020	0.00037	0.00077	0.00109	0.00020	0.00017	0.00024	0.00006
女性	0.00082	0.00167	0.00224	0.00197	0.00001	0.00003	0.00011	0.00002
城镇	0.00003	0.00007	0.00054	0.00024	0.00042	0.00018	0.00021	0.00013
农村	0.00002	0.00001	0.00007	0.00055	0.00100	0.00081	0.00090	0.00038
东部	0.00059	0.00109	0.00169	0.00095	0.00093	0.00054	0.00064	0.00023
中部	0.00026	0.00059	0.00092	0.00137	0.00096	0.00076	0.00099	0.00041
西部	0.00009	0.00007	0.00023	0.00084	0.00102	0.00117	0.00090	0.00056

续表

年份	教育的间接效应				社会经济地位(SES)的间接效应			
	2010	2012	2014	2017	2010	2012	2014	2017
50 后	0.00117	0.00156	0.00015	0.00238	0.00084	0.00053	0.00050	0.00043
60 后	0.00113	0.00157	0.00091	0.00184	0.00163	0.00126	0.00186	0.00076
70 后	0.00116	0.00145	0.00172	0.00209	0.00111	0.00097	0.00150	0.00079
80 后	0.00033	0.00072	0.00067	0.00088	0.00032	0.00031	0.00036	0.00026

资料来源：根据 CMDS 数据计算得到。

表 14 环境集对机会不平等的间接效应（相对值）

单位：%

年份	教育的间接效应				社会经济地位(SES)的间接效应			
	2010	2012	2014	2017	2010	2012	2014	2017
全样本	3.27	5.27	7.05	8.43	7.59	5.82	0.51	2.02
男性	2.96	5.24	11.42	11.73	2.96	2.41	2.52	0.60
女性	2.37	26.62	24.12	25.81	0.29	0.50	1.18	0.26
城镇	0.29	0.70	4.24	1.78	4.06	1.84	1.65	0.97
农村	0.18	0.10	0.60	3.71	8.77	7.53	7.30	2.56
东部	4.74	8.83	10.16	5.86	7.47	4.28	3.85	1.42
中部	2.09	4.68	5.68	7.80	7.73	6.03	6.11	2.33
西部	0.80	0.60	2.00	6.11	9.30	9.71	7.64	4.07
50 后	11.99	16.93	2.20	17.10	8.61	5.75	7.34	3.09
60 后	6.80	9.67	5.25	13.65	9.81	7.76	10.73	5.64
70 后	7.16	9.31	8.34	11.23	6.85	6.23	7.27	4.24
80 后	4.37	7.68	6.76	5.95	4.23	3.30	3.63	1.76

资料来源：根据 CMDS 数据计算得到。

通过对比表 11 和表 14，在全样本中，两种模型方法测算的总的间接效应的相对值相差不大。在 2017 年，基于式 (5) 和式 (6) 测算的 6.42% 的相对值略低于反事实模型的 10.45%，而测算的环境因素通过努力变量对机会不平等的间接影响的均值在 8.4% 左右，与反事实模型得出的 9.9% 接近，但在不同分组上其相对值有所差异。我们还可以看出，反事实模型在异质性分组测量教育的间接效应时，较好地纠正了前述模型测算带来的偏差，但在社会经济地位的间接效应的测算中，相较于反事实模型，前述模型的测算与其他相关学者得出的结论显得更合理一些。两种模型具有较好的互补性。

六 结论

本文依托 Roemer (1993) 提出的机会均等的“环境-努力”二元分析框架, 基于事前估计视角和参数法, 采用 2010 年至 2017 年中国流动人口动态监测调查 (CMDS) 数据, 测算了目前中国流动人口中机会不平等的程度及机会不平等在收入不平等中所占的比例。本文结合 Shapley 分解法计算出各个环境因素对机会不平等的贡献度, 并分别引入与环境相关的努力变量来测算环境对机会不平等的间接效应, 最后使用反事实模型对上述结果进行稳健性检验。本文的具体结论如下:

在机会不平等程度方面, 收入不平等和机会不平等从 2010 年至 2017 年均呈现逐年递增的趋势, 机会不平等占收入不平等的比重从 2010 年的 14% 增长到 2017 年的 16%。从机会不平等的相对水平上看, 农村高于城镇, 80 后小于 50 后、60 后和 70 后。正如上文提到的一样, 在存在未观测到的环境因素的情况下, 不论是使用参数法还是非参数法, 得到的都是对机会不平等的下限估计。在考虑到环境因素通过影响努力因素对机会不平等的间接效应后, 流动人口机会不平等的相对值在 22.38% ~ 24.61% 之间。

通过借助 Shapley 分解法分解环境因素对机会不平等的贡献, 我们发现, 不论是在总样本中还是在分组的样本中, 性别因素都占据着支配地位, 有别于相关研究认为户籍因素是目前中国收入不平等的主要因素的观点。同时, 我们也发现在性别分组中, 男性群体和女性群体在环境因素的贡献度上有着明显的差别。男性群体中年龄因素占据 70% 左右的比重, 而女性群体中户籍因素是主导因素, 这表明流动人口中的女性群体面临着来自性别和户籍的双重压力。

在测算环境对机会不平等的间接效应中, 通过分别引入与环境相关的努力变量, 我们发现, 教育间接效应的相对值逐年增加, 社会经济地位间接效应的相对值呈减少趋势。再次使用 Shapley 分解法分解环境因素对机会不平等的贡献, 教育变量的贡献度呈现逐步增强的趋势, 而社会经济地位的贡献度则逐年降低。在新构建的环境集中, 年龄、性别、户籍和出生地对机会不平等的总贡献度也从 2010 年的 47.24% 逐年增加至 2017 年的 56.89%。通过构建反事实模型对上述结果进行稳健性检验, 我们发现两种方法得出的结果近似, 并呈现相同的变化趋势。因此, 本文结果是稳健的。

参考文献：

- 董丽霞 (2018), 《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》, 《劳动经济研究》第 1 期, 第 44 - 62 页。
- 龚锋、李智、雷欣 (2017), 《努力对机会不平等的影响：测度与比较》, 《经济研究》第 3 期, 第 76 - 90 页。
- 呼倩、黄桂田 (2019), 《改革开放以来中国劳动力流动研究》, 《上海经济研究》第 6 期, 第 49 - 58 页。
- 江求川、任洁、张克中 (2014), 《中国城市居民机会不平等研究》, 《世界经济》第 4 期, 第 111 - 138 页。
- 雷欣、贾亚丽、龚锋 (2018), 《机会不平等的衡量：参数测度法的应用与改进》, 《统计研究》第 4 期, 第 73 - 85 页。
- 李莹 (2019), 《收入不平等变动的根源探析——基于机会不平等的测度》, 《云南财经大学学报》第 8 期, 第 12 - 23 页。
- 李莹、吕光明 (2018), 《我国城镇居民收入分配机会不平等因何而生》, 《统计研究》第 9 期, 第 67 - 78 页。
- 李莹、吕光明 (2019), 《中国机会不平等的生成源泉与作用渠道研究》, 《中国工业经济》第 9 期, 第 60 - 78 页。
- 刘波、胡宗义、龚志民 (2020), 《中国收入差距中的机会不平等再测度——基于“环境 - 能力 - 收入”的新思路》, 《南开经济研究》第 4 期, 第 107 - 126 页。
- 刘波、王修华、彭建刚 (2015), 《我国居民收入差距中的机会不平等——基于 CGSS 数据的实证研究》, 《上海经济研究》第 8 期, 第 77 - 88 页。
- 马占利、邹薇 (2018), 《中国机会不平等的测算与分解——基于“反事实”收入分布方法》, 《经济问题探索》第 11 期, 第 1 - 9 页。
- 史新杰、卫龙宝、方师乐、高叙文 (2018), 《中国收入分配中的机会不平等》, 《管理世界》第 3 期, 第 27 - 37 页。
- 叶文平、李新春、陈强远 (2018), 《流动人口对城市创业活跃度的影响：机制与证据》, 《经济研究》第 6 期, 第 157 - 170 页。
- Aaberge, Rolf, Magne Mogstad & Vito Peragine (2011). Measuring Long-term Inequality of Opportunity. *Journal of Public Economics*, 95 (3 - 4), 193 - 204.

- Betancort, Moises, Sara Darias-Curvo, Leopoldo Cabrera, Gustavo Marrero, Carmen Pérez, Juan Rodríguez & Daniel Sánchez (2019). Inequality of Opportunity in An Outermost Region: The Case of the Canary Islands. *Island Studies Journal*, 14 (2), 23 – 42.
- Bourguignon, François, Francisco Ferreira & Marta Menéndez (2007). Inequality of Opportunity in Brazil. *The Review of Income and Wealth*, 53 (4), 585 – 618.
- Brunori, Paolo, Vito Peragine & Laura Serlenga (2019). Upward and Downward Bias when Measuring Inequality of Opportunity. *Social Choice and Welfare*, 52 (4), 635 – 661.
- Checchi, Daniele & Vito Peragine (2010). Inequality of Opportunity in Italy. *The Journal of Economic Inequality*, 8 (4), 429 – 450.
- Davillas, Apostolos & Andrew Jones (2020). Ex Ante Inequality of Opportunity in Health, Decomposition and Distributional Analysis of Biomarkers. *Journal of Health Economics*, 69, 102251.
- Dias, Rosa (2009). Inequality of Opportunity in Health: Evidence from A UK Cohort Study. *Health Economics*, 18 (9), 1057 – 1074.
- Ferreira, Francisco & Jeremie Gignoux (2011). The Measurement of Inequality of Opportunity: Theory and an Application to Latin America. *The Review of Income and Wealth*, 57 (4), 622 – 657.
- Fleurbaey, Marc & Vito Peragine (2013). Ex Ante Versus Ex Post Equality of Opportunity. *Economica*, 80 (317), 118 – 130.
- Fleurbaey, Marc, Vito Peragine & Xavier Ramos (2017). Ex Post Inequality of Opportunity Comparisons. *Social Choice and Welfare*, 49 (3 – 4), 577 – 603.
- Foster, James & Artyom Shneyerov (2000). Path Independent Inequality Measures. *Journal of Economic Theory*, 91 (2), 199 – 222.
- Juárez, Florian & Isidro Soloaga (2014). Iop: Estimating Ex-Ante Inequality of Opportunity. *The Stata Journal*, 14 (4), 830 – 846.
- Lefranc, Arnaud, Nicolas Pistoiesib & Alain Trannoy (2009). Equality of Opportunity and Luck: Definitions and Testable Conditions, with An Application to Income in France. *Journal of Public Economics*, 93 (11 – 12), 1189 – 1207.
- Niehues, Judith & Andreas Peichl (2014). Upper Bounds of Inequality of Opportunity: Theory and Evidence for Germany and the US. *Social Choice and Welfare*, 43 (1), 73 – 99.

- Palomino, Juan, Gustavo Marrero & Juan Rodríguez (2019). Channels of Inequality of Opportunity: The Role of Education and Occupation in Europe. *Social Indicators Research*, 143 (3), 1045 – 1074.
- Roemer, John (1993). A Pragmatic Theory of Responsibility for the Egalitarian Planner. *Philosophy & Public Affairs*, 22 (2), 146 – 166.
- Roemer, John (1998). *Equality of Opportunity*. Cambridge: Harvard University Press.
- Roemer, John & Alain Trannoy (2016). Equality of Opportunity: Theory and Measurement. *Journal of Economic Literature*, 54 (4), 1288 – 1332.
- Trannoy, Alain, Sandy Tubeuf, Florence Jusot & Marion Devaux (2010). Inequality of Opportunities in Health in France: A First Pass. *Health Economics*, 19 (8), 921 – 938.

Estimating Opportunity Inequality among Migrants in China: An Ex-ante Perspective

Liu Lin & Li Meng

(School of Economics and Management, Shihezi University)

Abstract: In this paper we estimate the degree of opportunity inequality among migrants in China using the CMDS data and further analyze the contribution of different factors with Shapley decomposition. Our results show that from 2010 to 2017, both the income inequality index and opportunity inequality index of the migrants are on the increase. Opportunity inequality is more serious among rural households and in older age groups, and gender is the most important factor causing opportunity inequality. After the indirect effect of social environment on opportunity inequality being controlled, the relative contribution of opportunity inequality among migrants ranges from 22.38 percent to 24.61 percent. Finally, this paper performs a robustness check on the indirect effect of opportunity inequality with a counterfactual model.

Keywords: migrants, opportunity inequality, counterfactual model

JEL Classification: D63, J61, J71

(责任编辑: 合羽)