

# 农民工与城镇本地劳动力工资差距的再思考

## ——社会保险参保视角下的实证分析

孙婧芳

**摘要：**本文采用2023年中国社会综合调查数据，结合省级层面城镇职工社会保险缴费标准，以社会保险参保情况为基础对农民工和城镇本地劳动力工资进行调整，以期在初次分配维度上探讨农民工与城镇本地劳动力工资差距的来源。研究发现：农民工与城镇本地劳动力之间的工资差距主要源于个体特征差异，其中，受教育水平是工资差距的核心来源，其次是工作经验。经社会保险缴费调整得到的更接近初次分配口径的小时工资中，二者之间工资差距的来源依然以特征效应为主，但农民工获得的反歧视下降，工资歧视在一定程度上依然存在。在现行社会保险制度下，农民工参加城镇职工社会保险的比例较低可能是其自主选择的结果，其更加偏好当期收入。从年龄异质性来看，青年群体中农民工与城镇本地劳动力之间的工资差距主要来自特征效应，但中年农民工则面临工资歧视。鉴于此，应健全与常住人口相适应的教育资源配置机制，深化完善社会保障体系，优化社会保险制度设计，加强劳动者就业权益保障，以促进高质量充分就业并完善收入分配制度。

**关键词：**社会保险 初次分配收入 工资差距 农民工

**中图分类号：**F323.6；F249.24；C924.2 **文献标识码：**A

DOI:10.20074/j.cnki.11-3586/f.2026.03.011

### 一、引言

提高农民工参保率是“十五五”时期健全社会保障体系的重要目标之一。从社会保险制度的覆盖范围来看，随着《中华人民共和国社会保险法》的出台以及城乡居民

**[资助项目]** 中国社会科学院智库基础研究项目“促进高质量充分就业与提高劳动收入在初次分配中的比重”（编号：ZKJC250701）。

**[作者信息]** 孙婧芳，中国社会科学院经济研究所，电子邮箱：sunjf@cass.org.cn。

**[致谢]** 感谢匿名审稿专家的宝贵意见和建议，并感谢中国社会科学院大学经济学院硕士研究生朱尚格、蔡炎珊和赵欣楠在社会保险缴费信息收集方面提供的协助。

基本养老保险和基本医疗保险制度的全面建立，农民工在制度层面并未被排除在社会保险之外。然而，农民工参保率低这一现象持续存在。在一系列关于农民工市民化、基本公共服务均等化的研究中，农民工参加城镇职工社会保险的比例始终在15%~25%，没有出现明显提升（王春光，2006；李培林和李炜，2007；魏后凯和苏红键，2013；张培文等，2023）。结合农民工在劳动力市场中的表现，农民工被社会保险制度覆盖但实际参保率低的现状，可能源于两类作用机制截然不同的因素：其一，农民工在劳动力市场中受到歧视，其工资水平被压低，从而降低其参保能力；其二，农民工的人力资本得到充分实现，并未遭遇歧视，其不参保行为更可能是自主选择的结果。这两类因素所对应的提高农民工参保率的政策具有显著差异。如果农民工是遭遇歧视导致工资水平偏低而未能参保，则政策侧重点应放在劳动力市场的改革上。随着劳动力市场歧视问题的逐步缓解，当农民工获得较为公平的工资分配之后，其参保率自然会有所提高。如果农民工并未遭遇工资歧视，而是出于自主选择不参保，则政策应侧重于完善社会保险制度，改善农民工参保环境，从而提高其参保率。鉴于此，本文将探讨农民工在劳动力市场初次分配时所面临的工资歧视问题，进而明晰中国健全社会保险制度、提高农民工参保率的微观基础，以期增强相关政策的靶向性与实施效能。

自2004年前后东部沿海地区出现“民工荒”以来，农民工工资明显上涨。学术界从不同角度对该问题展开讨论。在宏观层面，从人口结构变迁出发，蔡昉（2010）认为，中国的“刘易斯转折点”已经出现，劳动力市场供需结构的变化推动农民工工资上涨。在微观层面，围绕劳动力市场分割视角，诸多学者运用各项微观调查数据和分析方法，剖析了农民工与城镇本地劳动力工资差距及其成因。近年来多数研究表明，尽管农民工工资仍低于城镇本地劳动力，但二者之间的工资差距已显著缩小，且这种差距主要源于个体特征差异，歧视成分较低，甚至出现了反歧视现象（陈昊等，2017；章莉和吴彬彬，2019；江求川和鲁元平，2024）。然而，早期研究普遍发现，农民工与城镇本地劳动力之间的工资差距较大，且歧视是造成差距的主要原因（王美艳，2005；邓曲恒，2007；余向华和陈雪娟，2012）。针对农民工与城镇本地劳动力工资差距成因的上述阶段性变化，一些学者利用长期微观调查的多轮数据，在统一的分析方法和数据结构下，系统考察了二者之间工资差距的动态演变过程和来源。研究表明，农民工与城镇本地劳动力的工资呈现显著的趋同态势，无论是相对工资的变化趋势，还是不同时期相对工资水平的变化特征，都呈现趋同的特点，而且，二者之间工资差距的来源也逐步显现反歧视倾向（Zhang，2020；邢春冰等，2021）。

然而，受访者在接受问卷调查时，基于理解和填写习惯，其填报的工资多为实发工资，而非初次分配的市场化收入。即便进一步追问受访者，也难以准确获取其社会保险缴纳金额、个人所得税代扣代缴额度等关键信息。此外，值得注意的是，社会保

险缴费中企业承担部分远高于个人缴费部分，而个人在核算自身收入时，通常并未将企业代缴的社会保险纳入统计范围。已有研究往往忽视了微观调查中这一潜在的收入核算偏差，直接将问卷中受访者填报的工资作为初次分配收入展开分析，这在一定程度上削弱了研究结论的准确性。为了更加深入地探析劳动力市场运行规律，以及劳动者在当期收入与社会保险参保或退休收入之间的选择偏好特征，亟须立足社会保险参保视角，对农民工与城镇本地劳动力之间的工资差距展开再思考，厘清其中的微观事实<sup>①</sup>。若经社会保险缴费调整后，农民工与城镇本地劳动力之间的工资差距由个体特征差异所主导，歧视的影响占比很低甚至存在反歧视，则农民工群体偏低的社会保险参保率更可能是其在现行社会保险制度下自主选择的结果。反之，若调整之后，农民工与城镇本地劳动力之间的工资差距中歧视的影响所占比例大幅提高，则表明农民工在劳动力市场中遭遇的歧视并未消除，由此导致的参保能力弱化亦是其参保率偏低的重要原因。如前文所述，这两种情况对应的政策优化方向存在明显差异。

本文采用2023年中国社会状况综合调查数据，结合各省份社会保险缴费政策，考察农民工与城镇本地劳动力经社会保险缴费调整后在初次分配阶段的工资差距。通过系统梳理省级层面的社会保险缴费标准，并将其与微观调查数据进行匹配，本文构建了包含受访者工资性收入、社会保险缴费水平以及个体特征等信息的微观数据，为更精准地测度农民工与城镇本地劳动力在初次分配阶段的工资差距提供了翔实可靠的数据支撑。本文的主要贡献在于：其一，针对问卷调查中因受访者理解和填写习惯可能产生的偏差，对受访者填报的工资进行核算与调整，使其更接近于劳动力市场初次分配的结果，从而更有效地反映劳动力市场的市场化程度，进而为相关政策建议奠定坚实的微观基础。其二，运用无条件分位数回归方法对调整后的初次分配收入展开实证分析，以充分验证劳动力市场的工资分布特征。其三，基于对初次分配的分析结果，进一步探讨劳动者在当期收入与退休收入之间的选择偏好。

## 二、从社会保险参保视角分析工资差距的必要性

### （一）社会保险缴纳与工资统计偏差

非农就业劳动力内部的社会保险制度安排存在较大差异。农民工群体既被城镇

<sup>①</sup> 国家税务总局数据显示，2023年，“我国取得综合所得的人员中，无需缴纳个税的人员占比超过七成，在剩余不到三成的实际缴税人员中，60%以上仅适用3%的最低档税率，缴税金额较少……申报收入位居全国前10%的个人缴纳的个税占全部个税的九成以上”（资料来源：《2023个税汇算清缴数据发布 超七成人员无需缴纳个税》，<https://www.chinatax.gov.cn/chinatax/n810219/n810780/c5235245/content.html>）。鉴于无需缴纳个人所得税的人员占比较高，且最高收入群体是个人所得税主要贡献者，本文在折算受访者初次分配收入时，仅考虑其社会保险的缴纳情况，忽略其个人所得税的缴纳情况。

职工社会保险体系覆盖，亦被城乡居民社会保险体系覆盖。因此，如果农民工参加城乡居民社会保险，城镇本地劳动力参加城镇职工社会保险，而城镇本地劳动力在接受微观调查时并未清晰区分税前收入与单位缴纳的社会保险费用，其填报的工资往往是实发工资，那么从初次分配的角度来看，这会造成统计遗漏。尤其是城镇职工基本养老保险中企业缴费占比较高，如果遗漏了社会保险部分，将会给初次分配工资核算带来不容忽视的系统性偏差。2019年之前，城镇职工基本养老保险的企业缴费比例一般为本企业工资总额的20%。2019年《国务院办公厅关于印发降低社会保险费率综合方案的通知》出台后，该比例下调至16%。除城镇职工基本养老保险外，城镇职工基本医疗保险的企业缴费比例一般为本企业工资总额的8%左右。此外，企业还缴纳生育保险、失业保险和工伤保险。2019年之后，企业为职工缴纳的各类社会保险合计占工资总额的比例达25.7%~27.4%。由此可见，如果遗漏企业缴纳的社会保险部分，工资性收入的统计遗漏占比将高达约1/4；如果进一步遗漏个人缴费部分，仅以实发工资进行核算，统计遗漏占比会更高，接近40%<sup>①</sup>。与之不同的是，城乡居民社会保险由个人自行缴费，并不会从工资中直接扣除，也无企业缴纳部分。这意味着，参加城乡居民社会保险的劳动者在微观调查中填报的工资，本身即为初次分配收入。因此，受访者在填写调查问卷时，如果遗漏了社会保险部分，没有将其纳入收入核算，将导致参加城镇职工社会保险的劳动者与参加城乡居民社会保险的劳动者工资性收入数据缺乏可比性——前者填报的是税后实发工资，后者填报的是税前初次分配收入。

从当期收入与退休收入的权衡视角来看，社会保险制度（尤其是养老保险制度）的核心目标，在于帮助家庭及个人在整个生命周期内合理分配资源，将在职阶段的收入转移到退休时期（Ando and Modigliani, 1963）。这意味着，养老金的公平性在很大程度上是在职收入公平性的延续。劳动力市场的工资分配机制不仅直接影响劳动者当期的就业公平、工资公平等问题，而且关系到劳动者退休之后的收入公平性，劳动力市场公平性的影响贯穿于劳动者整个生命周期。从中国社会保险制度的实践来看，城乡居民社会保险缴费水平偏低的特征，在一定程度上反映出参加城乡居民社会保险的群体更偏重当期收入所得，相对而言，养老金收入处于次要地位。对于农民工群体而言，这种对当期收入的偏好可能与其返乡养老的意愿密切相关。农村地区的养老成本低于城镇地区，且城乡居民社会保险已经在一定程度上弱化了农村居民对依靠土地养

<sup>①</sup> 城镇职工社会保险的缴纳范围通常将基本养老保险、基本医疗保险、失业保险与工伤保险一并纳入，且从工资性收入角度来看，上述各项社会保险缴费均为工资性收入的重要组成部分。鉴于此，本文结合社会保险缴费情况调整劳动者工资性收入时，均以社会保险缴费总额为核算依据。

老的需求，并推动了农村土地流转（徐志刚等，2018；Zhu et al., 2022；洪甘霖等，2023）。从劳动力市场的分析视角出发，在探讨农民工与城镇本地劳动力的工资差距及劳动力市场中的歧视问题时，应充分考虑社会保险制度的影响，在相同的工资范畴内比较二者之间的工资差距并识别工资歧视，才能避免因统计口径不一致或统计遗漏产生的偏误，从而为相关政策的制定提供清晰准确的微观基础。

## （二）社会保险再分配功能与在职收入公平性的传递效应

社会保险具有重要的再分配功能，其作用效果远超个人所得税（岳希明和徐静，2024）。1994年，世界银行发布了《防止老龄危机——保护老年人及促进增长的政策》，针对人口老龄化问题提出了养老保险制度建设的基本原则。这些原则突出了养老保险的再分配功能，要求养老保险制度体系为老年人构筑安全保障网，帮助他们应对退出劳动岗位、遭遇意外变故以及早年不可逆决策所带来的困境，同时，通过激励就业与储蓄行为促进经济发展。此外，世界银行还提出养老保险制度“三大支柱”体系<sup>①</sup>。2005年，世界银行将养老保险制度进一步扩充至“五大支柱”体系，新增了非缴费型养老支柱以提供最低生活保障，并将家庭成员间或代际非正规保障纳入其中<sup>②</sup>。根据世界银行提出的养老支柱体系，为应对信息不对称和中青年时期决策偏差可能导致的老年困境，其中的两大支柱被明确设定为强制性制度安排，并强调“防贫兜底”和“保值增值”这两大核心目标。这种强制性设计从侧面反映出，如果完全依赖个人自主权衡当期收入与退休收入，其可能更倾向于选择当期收入，这不仅可能使个人和家庭陷入养老困境，还可能影响国家整体养老体系的可持续性。

从中国养老保险体系来看，其核心框架由三大支柱构成：第一支柱是基本养老保险，涵盖城镇职工基本养老保险和城乡居民基本养老保险两大板块；第二支柱是企业年金和职业年金；第三支柱是个人自愿性养老保险，包括个人储蓄性养老保险和商业养老保险。其中，第一支柱是中国养老保险体系的主体部分。随着城乡居民基本养老保险的全面推行与持续深化，中国养老保险的覆盖率大幅提升。2024年末，中国16岁及以上人口规模为11.68亿人，其中，参加基本养老保险的人口为10.73亿人，参保率超过90%，如果剔除在校学生群体，参保率将进一步提高；从就业人员维度来看，

① World Bank, 1994, "Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth", <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/973571468174557899/averting-the-old-age-crisis-policies-to-protect-the-old-and-promote-growth>.

② World Bank, 2005, "Old Age Income Support in the 21st Century: An International Perspective on Pension Systems and Reform", <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/363541468142778753>.

参加城镇职工基本养老保险的比例为52.71%<sup>①</sup>。这意味着，中国养老保险第一支柱内部存在较大差异。已有研究表明，中国养老保险体系呈现“多轨制”特征，具有较强的分割性，城乡差异、群体差异较为明显（李实等，2013；朱玲，2014）。此外，养老金收入的不平等不仅表现为制度内部的差异，还体现在不同制度之间的差异上，其中，制度之间的差异是养老金收入不平等的主要来源（贾晗睿等，2021）。

由此可见，社会保险的再分配功能和养老金的公平性问题在全球范围内都受到广泛关注。如前文所述，养老金与劳动者在职期间的工资性收入密切相关，在一定程度上也反映了劳动者在职期间工资性收入的公平性特征。如果劳动者在职期间的工资性收入是公平的，那么养老金所呈现的不平等或再分配效应不足，更可能源于养老金制度本身；反之，如果劳动者在职期间的工资性收入存在非公平性，那么养老金面临的问题不仅仅是制度性的，还与劳动力市场的特征息息相关。因此，综合考虑社会保险制度对工资差距的影响，有助于更加清晰地理解劳动力市场的特征，并从劳动力市场视角为进一步完善养老保险等社会保险制度提供决策依据。

### 三、研究设计

#### （一）数据来源

本文使用的数据来自2023年中国社会状况综合调查（Chinese Social Survey, CSS）数据库。CSS是中国社会科学院社会学研究所于2005年发起的一项全国范围内的大型连续性抽样调查项目，属于双年度的纵贯调查。从调查区域来看，2023年，CSS覆盖全国31个省份624个村居，具有较强的全国代表性。从调查内容来看，CSS涵盖人口特征、就业特征、收入及其来源、社会保险等多个维度，这为本文的研究提供了较为完备的数据基础。

本文研究的核心问题是劳动力市场中的工资差距及其来源，并在此基础上分析劳动者的社会保险参保行为。聚焦于“十五五”时期提高农民工参保率这一民生目标，本文以在城镇劳动力市场中就业的农民工和城镇本地劳动力为主要研究对象。根据2023年CSS问卷中关于户口信息的调查，将农业户籍以及在获得居民户口之前为农业户籍的群体界定为农业户籍人口。结合就业特征，将非农就业的农业户籍人口识别为农民工，将非农户籍人口根据其户籍所在地是否为调查时的工作地识别出是否为城镇

<sup>①</sup> 资料来源：《中华人民共和国2024年国民经济和社会发展统计公报》，[https://www.stats.gov.cn/sj/zxfb/202502/t20250228\\_1958817.html](https://www.stats.gov.cn/sj/zxfb/202502/t20250228_1958817.html)；《2024年度人力资源和社会保障事业发展统计公报》，[https://www.mohrss.gov.cn/xxgk2020/fdzd/gknr/ghtj/tj/ndtj/202506/t20250616\\_543689.html](https://www.mohrss.gov.cn/xxgk2020/fdzd/gknr/ghtj/tj/ndtj/202506/t20250616_543689.html)。

本地劳动力。在对数据进行初步筛选的基础上，本文保留18~64岁的非农就业劳动力样本，剔除存在变量缺失的样本，并对关键变量进行1%的缩尾处理。经数据清理与筛选后，有效样本包含农民工1966个、城镇本地劳动力1125个<sup>①</sup>。

此外，本文还收集整理了2023年各省份的社会保险缴费标准以及社会保险缴费基数上下限。根据受访者的工作地，将个体层面的微观数据与省级层面的社会保险缴费情况进行匹配，最终得到本文使用的基础数据<sup>②</sup>。

## （二）社会保险参保情况

依据农民工和城镇本地劳动力是否参加城镇职工社会保险对其工资进行调整，是本文后续分析的重要基础。因此，有必要在此对两类群体的参保情况进行描述性说明<sup>③</sup>。

由表1可知，城镇本地劳动力参加城镇职工基本养老保险、基本医疗保险、失业保险和工伤保险的比例均在70%以上，其中，城镇职工基本养老保险的参保比例最高，达到79.52%。相较之下，农民工参加城镇职工社会保险的比例远低于城镇本地劳动力，其中，工伤保险的参保比例最高，达到44.00%。从当期收入与退休收入的权衡视角来看，农民工更偏好当期收入。这一倾向表现为农民工养老保险参保率整体偏低。具体来说，城镇本地劳动力中既没有参加城镇职工基本养老保险，也没有参加城乡居民基本养老保险的比例仅为11.90%，而农民工中这一比例则达到30.70%。此外，农民工参加城镇职工基本养老保险的比例明显低于城镇本地劳动力。在已参加养老保险的群体中，农民工参加城镇职工基本养老保险的比例仅为53.46%，而城镇本地劳动力相应比例高达90.25%。将两类群体按工资水平划分为低工资（0~25分位）、中低工资（25~50分位）、中高工资（50~75分位）和高工资（75~100分位）四组，对比其参保分布可以发现：无论是城镇本地劳动力还是农民工，参加城镇职工社会保险的比例均随工资水平的提高而上升，不仅未参保比例下降，而且参加城乡居民社会保险的比例也呈下降趋

① 本文仅选取农民工与城镇本地劳动力作为分析对象，主要原因在于二者在农业户籍与非农业户籍属性、个体特征等方面差异显著，且农民工工资水平长期处于相对低位。与由乡到城流动的劳动力不同，城市之间流动的劳动力的人力资本水平往往高于城镇本地劳动力，且其工资性收入也处于较高水平。

② 由于微观调查的限制，2023年CSS数据仅能识别至省级行政区划代码，地市级编码为自由编码，并非可识别的国标编码。虽然难以识别至地市级，但这一限制对社会保险缴费比例的界定影响较小。通过整理城镇职工社会保险缴费比例可以发现，同一省份内部各地市之间的缴费比例的差异并不明显，甚至各省份之间缴费比例的差异也较小，省际差异主要体现在缴费基数上。即便各省份内部不同地市之间的缴费基数存在一定差异，该差异通常也小于省际差异。因此，在采用省级层面的社会保险缴费比例和缴费基数对劳动者初次分配收入进行推算处理时，所得结果依然具备较高的准确性。

③ 本文研究样本仅限于受雇劳动力，未包括企业经营者和自营劳动者。从劳动力市场公平性视角来看，企业经营者和自营劳动者的收入除劳动报酬外，还包含资本收益，其收入状况难以直接和清晰地反映劳动力市场的分配结果。

势。这表明，收入水平对劳动者参加城镇职工社会保险具有重要影响。因此，工资是否充分反映劳动者人力资本水平尤为值得关注，其背后蕴含着重要的政策含义。

表1 农民工和城镇本地劳动力社会保险参保分布特征 单位:%

项目	养老保险			医疗保险			失业保险		工伤保险	
	未参保	职工类	居民类	未参保	职工类	居民类	未参保	职工类	未参保	职工类
城镇本地劳动力	11.90	79.52	8.59	5.17	75.11	19.71	28.20	71.80	25.97	74.03
0~25分位	22.77	63.37	13.86	10.84	57.64	31.53	54.19	45.81	51.74	48.26
25~50分位	14.52	72.20	13.28	5.44	69.87	24.69	34.17	65.83	31.38	68.62
50~75分位	9.82	86.18	4.00	5.04	80.22	14.75	22.10	77.90	18.89	81.11
75~100分位	6.36	87.53	6.11	2.29	83.46	14.25	15.38	84.62	14.21	85.79
农民工	30.70	37.05	32.25	10.52	34.35	55.13	62.65	37.35	56.00	44.00
0~25分位	38.26	17.44	44.31	12.41	14.54	73.05	84.68	15.32	80.70	19.30
25~50分位	32.70	36.48	30.81	13.07	33.52	53.41	62.08	37.92	53.93	46.07
50~75分位	26.44	46.91	26.65	8.74	44.56	46.70	51.69	48.31	44.09	55.91
75~100分位	21.37	55.89	22.74	6.58	53.70	39.73	42.62	57.38	35.07	64.93

结合农民工和城镇本地劳动力在参保结构上的差异，可进一步凸显在工资差距分析中纳入社会保险缴费因素的重要性。若忽视社会保险缴费因素，尤其是企业缴费部分，不仅难以准确揭示劳动力市场中工资歧视的真实特征，也难以有效识别农民工在当期收入与退休收入之间的权衡是否为自主选择的结果。

### （三）变量选择

1. 被解释变量。本文以小时工资的对数作为被解释变量。具体而言，本文将劳动者主要非农工作的月工资（以下简称“问卷月工资”）除以月工作小时数，计算得出问卷小时工资<sup>①</sup>。基于微观调查数据，本文尝试通过社会保险参保情况倒推劳动者在劳动力市场中初次分配的工资性收入。由于城乡居民社会保险由参保人自行缴纳，对于参加城乡居民社会保险的劳动者而言，其调查中填报的工资可视为其初次分配收入所得，无需根据参保情况对工资性收入进行调整。与之不同的是，参加城镇职工社会保险的劳动者，其社会保险缴费不仅包含企业缴纳部分，而且个人缴纳部分也由企业代扣代缴。鉴于此，本文仅针对参加城镇职工社会保险的劳动者的问卷工资进行相应调整。

根据城镇职工社会保险缴费标准，本文将其与样本数据进行匹配时，分三种情境进行测算（见表2）。首先，情境I假设所有参加城镇职工社会保险的劳动者均按照社会保险缴费基数下限进行缴费，经社会保险缴费调整后的工资简称“调整月工资I”。

<sup>①</sup> 如果受访者从事多份工作，则文中统计的月工资将会低于受访者的实际月收入。根据问卷中关于非农工作份数的调查结果，在全部受访者中，仅从事1份非农工作的受访者占比为95.44%；经样本筛选后，本文使用的样本中，仅从事1份非农工作的受访者占比为95.88%。本文重点分析劳动者的小时工资，一方面可以缓解从事多份非农工作导致的所统计主要非农工作的月工资低于实际月收入的偏差，另一方面也能够避免工作时长差异对月工资的影响。

其次，情境Ⅱ假设按照劳动者实际收入缴纳城镇职工社会保险，经社会保险缴费调整后的工资简称“调整月工资Ⅱ”。具体而言，如果月工资<sup>①</sup>高于社会保险缴费基数下限且低于上限，则调整月工资Ⅱ的计算公式为：调整月工资Ⅱ=月工资×（1+就业地企业缴费比例）。如果月工资低于社会保险缴费基数下限，则计算公式为：调整月工资Ⅱ=问卷月工资+社会保险缴费基数下限×（就业地个人缴费比例+就业地企业缴费比例）。如果月工资高于社会保险缴费基数上限，则计算公式为：调整月工资Ⅱ=问卷月工资+社会保险缴费基数上限×（就业地个人缴费比例+就业地企业缴费比例）。最后，情境Ⅲ将问卷月工资视为劳动者的应税收入，假设按照劳动者应税收入缴纳城镇职工社会保险，经社会保险缴费调整后的工资简称“调整月工资Ⅲ”。与情境Ⅱ类似，将问卷月工资（应税收入）与社会保险缴费基数上下限进行比较，分为三种情况分别结合就业地企业缴费比例测算得到调整月工资Ⅲ。

表2 不同社会保险缴费情境下初次分配月工资

情境Ⅰ (按照社会保险缴费基数下限调整)	情境Ⅱ (按照实际收入调整)	情境Ⅲ (按照应税收入调整)
调整月工资Ⅰ=问卷月工资+社会保险缴费基数下限×(就业地个人缴费比例+就业地企业缴费比例)	①月工资≤社会保险缴费基数下限： 调整月工资Ⅱ=问卷月工资+社会保险缴费基数下限×(就业地个人缴费比例+就业地企业缴费比例)	①问卷月工资≤社会保险缴费基数下限： 调整月工资Ⅲ=问卷月工资+(社会保险缴费基数下限×就业地企业缴费比例)
	②社会保险缴费基数下限<月工资≤社会保险缴费基数上限： 月工资=问卷月工资/(1-就业地个人缴费比例) 调整月工资Ⅱ=月工资×(1+就业地企业缴费比例)	②社会保险缴费基数下限<问卷月工资≤社会保险缴费基数上限： 调整月工资Ⅲ=问卷月工资×(1+就业地企业缴费比例)
	③月工资>社会保险缴费基数上限： 调整月工资Ⅱ=问卷月工资+社会保险缴费基数上限×(就业地个人缴费比例+就业地企业缴费比例)	③问卷月工资>社会保险缴费基数上限： 调整月工资Ⅲ=问卷月工资+(社会保险缴费基数上限×就业地企业缴费比例)

通过整合社会保险缴费情况，本文测算了农民工和城镇本地劳动力在劳动力市场中获得的初次分配收入（见表3）。由表3可知，农民工与城镇本地劳动力的月工资和小时工资都存在一定差距，城镇本地劳动力的工资水平明显高于农民工。根据问卷调查中直接获得的工资性收入数据，城镇本地劳动力的月工资和小时工资分别是农民工的1.30倍和1.39倍。经社会保险缴费调整后，两类群体之间的工资差距进一步扩大。以社会保险缴费基数下限测算的调整工资Ⅰ为例，城镇本地劳动力与农民工在月工资和小时工资上的差距，分别从问卷工资的1.30倍和1.39倍扩大至调整工资Ⅰ的1.39倍

① 月工资=问卷月工资/(1-就业地个人缴费比例)，将测算得到的劳动者月工资与当地社会保险缴费基数上下限进行比较，据此判断其缴费水平。

和1.50倍。调整工资Ⅱ和调整工资Ⅲ所反映的城镇本地劳动力与农民工之间的工资差距也呈现相同的趋势。城镇本地劳动力的工资水平较高，且在将社会保险缴费纳入工资核算之后，进一步扩大了城镇本地劳动力与农民工之间的工资差距。此外，值得注意的是，城镇本地劳动力与农民工之间的月工资差距小于小时工资差距，这表明农民工的月工作时长普遍更长，劳动强度较大。

**表3 不同社会保险缴费情境下农民工和城镇本地劳动力工资分布情况** 单位:元

项目	农民工				城镇本地劳动力			
	月工资		小时工资		月工资		小时工资	
	均值	标准差	均值	标准差	均值	标准差	均值	标准差
问卷工资 <sup>a</sup>	5071.191	4497.677	28.660	32.364	6607.617	5555.565	39.963	42.392
调整工资Ⅰ <sup>b</sup>	5628.561	4750.709	32.102	36.291	7837.870	5781.727	48.122	48.691
调整工资Ⅱ <sup>b</sup>	6000.008	5558.001	34.150	40.376	8834.720	7257.193	53.672	55.046
调整工资Ⅲ <sup>b</sup>	5674.353	5167.187	32.232	37.328	8050.749	6650.811	48.859	50.139
观测值 <sup>c</sup>	1966		1952		1125		1117	

注：a 问卷工资指未经社会保险缴费调整，通过问卷调查直接获得的工资性收入；b 调整工资Ⅰ、调整工资Ⅱ、调整工资Ⅲ分别为根据情境Ⅰ、情境Ⅱ、情境Ⅲ调整后得到的社会保险缴费前的工资性收入；c 月工资与小时工资的观测值存在略微差异，系少量样本缺失月工作时长数据所致。

2. 解释变量。本文根据 Mincer 工资方程设置解释变量。具体包括：劳动者的受教育年限、工作经验及其平方项、性别、婚姻状况和健康状况，并控制了省份和行业变量，以此来控制个体特征等因素对劳动者工资差距的影响。

变量定义及描述性统计结果见表4。

**表4 变量定义及描述性统计结果**

变量名称	变量定义	观测值	均值	标准差	最小值	最大值
<b>被解释变量</b>						
问卷小时工资	问卷月工资与月工作小时数之比（元）	3069	32.774	36.734	2.111	333.333
调整小时工资Ⅰ	调整月工资Ⅰ与月工作小时数之比（元）	3069	37.933	41.945	2.375	375.000
调整小时工资Ⅱ	调整月工资Ⅱ与月工作小时数之比（元）	3069	41.255	47.192	2.375	389.179
调整小时工资Ⅲ	调整月工资Ⅲ与月工作小时数之比（元）	3069	38.284	43.181	2.375	361.770
<b>解释变量</b>						
受教育年限	个体受教育年限（年）	3091	11.872	3.890	0.000	19.000
工作经验	个体当前主要工作的工作年限（年）	3091	17.085	12.110	0.000	49.000
性别	个体性别：男性=1，女性=0	3091	0.535	0.499	0	1
婚姻状况	个体是否已婚：是=1，否=0	3091	0.777	0.416	0	1
健康状况	个体身体健康状况：较差=1，其他=0	3091	0.047	0.211	0	1
	一般=1，其他=0	3091	0.379	0.485	0	1
	较好=1，其他=0	3091	0.574	0.495	0	1

注：①表中报告的是被解释变量原值统计结果，在回归分析中，被解释变量作取对数处理；②对问卷月工资、问卷小时工资进行1%的缩尾处理，调整小时工资依据缩尾处理后的问卷小时工资计算得到。

#### (四) 模型设定

本文根据 Mincer 工资方程，运用无条件分位数回归 (unconditional quantile regressions, UQR) 对农民工和城镇本地劳动力的小时工资进行估计。UQR 以影响函数 (influence function) 为基础，通过影响函数重新构造被解释变量，进而对构造的新被解释变量进行回归分析 (Firpo et al., 2009)。以 UQR 为基础的工资差距分解方法被称为 FFL 分解 (Firpo et al., 2018; Rios-Avila, 2020)。本文对 UQR 和 FFL 分解方法作如下说明。

基于被解释变量工资  $y$  的影响函数设定如下：

$$\text{IF}[y, v(F_Y)] = \lim_{t \rightarrow 0} \left[ \frac{v(F_{t, \Delta Y}) - v(F_Y)}{t} \right] = \left. \frac{\partial v(F_{t, \Delta Y})}{\partial t} \right|_{t=0} \quad (1)$$

构造新的被解释变量 RIF：

$$\text{RIF}[y, v(F_Y)] = v(F_Y) + \text{IF}[y, v(F_Y)] \quad (2)$$

(1) 式和 (2) 式中： $y$  表示劳动者的小时工资， $v(F_Y)$  表示  $y$  的分布， $\text{IF}[y, v(F_Y)]$  表示  $y$  的影响函数。

对新构造的被解释变量 RIF 进行最小二乘估计：

$$E\{\text{RIF}[y, v(F_Y)] | \mathbf{X}\} = \mathbf{X}\boldsymbol{\beta} \quad (3)$$

(3) 式中： $\mathbf{X}$  表示解释变量向量，包含受教育年限、工作经验及其平方项、性别、婚姻状况和健康状况，并控制了省份固定效应和行业固定效应。 $\boldsymbol{\beta}$  表示解释变量  $\mathbf{X}$  对被解释变量 RIF 的无条件边际影响。

以重加权方式将工资差距分解为结构效应和特征效应，进而对获得的实际收入分布和反事实收入分布进行无条件分位数估计，并通过 Oaxaca-Blinder 分解方法，将工资差距分解至各解释变量。

具体而言， $T$  是群体分类变量，其中， $T=1$  表示农民工， $T=0$  表示城镇本地劳动力。农民工和城镇本地劳动力的权重  $\hat{\omega}_1(T)$  和  $\hat{\omega}_0(T)$  可直接由样本分布得到：

$$\hat{\omega}_1(T) = \frac{T}{p} \quad (4)$$

$$\hat{\omega}_0(T) = \frac{1-T}{1-p} \quad (5)$$

基于倾向得分估计，构造反事实估计的权重  $\hat{\omega}_c(T, \mathbf{X})$ ：

$$\hat{\omega}_c(T, \mathbf{X}) = \frac{1-T}{p} \cdot \frac{\hat{p}(T=1 | \mathbf{X})}{1 - \hat{p}(T=1 | \mathbf{X})} \quad (6)$$

(4) 式、(5) 式和 (6) 式中： $p = N^{-1} \sum_{i=1}^N T_i$  是农民工样本数占总样本数的比例。 $\hat{p}(T = 1 | X)$  是在解释变量  $X$  给定的情况下，样本属于农民工群体的条件概率，由 Logit 模型估计得到。

基于加权核密度估计和重置的反事实权重，可得到城镇本地劳动力反事实的工资分布  $v_c$ 。具体步骤为：先将农民工样本与城镇本地劳动力样本合并，利用 Logit 模型估计倾向得分  $\hat{p}(T = 1 | X)$ ，再根据 (6) 式计算反事实权重，进而通过加权核密度估计得到城镇本地劳动力的反事实工资分布。该结果反映的是，在给定城镇本地劳动力工资结构的情况下，通过权重  $\hat{\omega}_c(T, X)$  调整城镇本地劳动力个体特征，使其趋近于农民工特征，从而模拟出如果城镇本地劳动力拥有与农民工相似的个体特征时，其所呈现的工资分布。

对工资差距进行分解：

$$\Delta v = v_1 - v_0 = (v_1 - v_c) + (v_c - v_0) = \Delta v_s + \Delta v_x \quad (7)$$

(7) 式中： $v_1$  表示农民工工资分布， $v_0$  表示城镇本地劳动力工资分布， $v_c$  表示城镇本地劳动力反事实的工资分布， $\Delta v$  表示农民工与城镇本地劳动力之间的工资总差距； $\Delta v_s$  表示结构效应，反映由身份差异导致的劳动回报率不同所形成的工资差距，一般被视为工资歧视； $\Delta v_x$  表示特征效应，即由不同群体劳动者个体特征差异所决定的工资差距，是工资差距中可解释的部分。该分解涉及基准组选择问题：理论上应选择无歧视的劳动力市场工资结构作为基准组，但实证研究中关于基准组的选择并未形成定论，通常根据劳动力市场实际情况确定。在农民工与城镇本地劳动力工资差距的研究中，一般以城镇本地劳动力的估计参数为基准。

进一步，通过最小二乘法估计城镇本地劳动力的反事实工资分布：

$$v_c = v(F_Y^c) = E\left\{RIF\left[y, v(F_Y^c)\right]\right\} = X_c \hat{\beta}_c \quad (8)$$

(8) 式中： $v(F_Y^c)$  表示反事实的工资分布， $X_c$  表示反事实的解释变量向量， $\hat{\beta}_c$  是其对应的估计参数。

基于 RIF 回归模型，可直接估计解释变量对工资的边际影响，并将不同群体之间的工资差距分解至各解释变量：

$$\begin{aligned} \Delta v &= v_1 - v_0 = (v_1 - v_c) + (v_c - v_0) = \Delta v_s + \Delta v_x \\ &= \left[ E(X_1) \hat{\beta}_1 - E(X_1) \hat{\beta}_c \right] + \left[ E(X_1) \hat{\beta}_c - E(X_0) \hat{\beta}_0 \right] \\ &= E(X_1) (\hat{\beta}_1 - \hat{\beta}_c) + \left[ E(X_1) - E(X_c) \right] \hat{\beta}_c \\ &\quad + \left[ E(X_c) - E(X_0) \right] \hat{\beta}_0 + E(X_c) (\hat{\beta}_c - \hat{\beta}_0) \end{aligned} \quad (9)$$

(9) 式中： $E(X_1)$ 、 $E(X_0)$ 、 $E(X_c)$ 分别表示农民工、城镇本地劳动力、城镇本地劳动力反事实的解释变量向量的期望， $\hat{\beta}_1$ 、 $\hat{\beta}_0$ 、 $\hat{\beta}_c$ 分别表示其对应的RIF回归的估计参数， $E(X_1)(\hat{\beta}_1 - \hat{\beta}_c)$ 和 $[E(X_1) - E(X_c)]\hat{\beta}_c$ 之和为结构效应， $[E(X_c) - E(X_0)]\hat{\beta}_0$ 和 $E(X_c)(\hat{\beta}_c - \hat{\beta}_0)$ 之和为特征效应，其中， $E(X_1)(\hat{\beta}_1 - \hat{\beta}_c)$ 和 $[E(X_c) - E(X_0)]\hat{\beta}_0$ 分别被视为纯结构效应和纯特征效应。本文重点分析结构效应、特征效应，以及纯特征效应中各解释变量的贡献。

#### 四、不同社会保险缴费情境下教育回报与工资歧视

##### (一) 不同社会保险缴费情境下教育回报的差异分析

在运用UQR时，本文在行业层面采用聚类稳健标准误进行估计。基于UQR回归结果，本文进一步分析不同社会保险缴费情境下小时工资的教育回报率（如图1所示）。

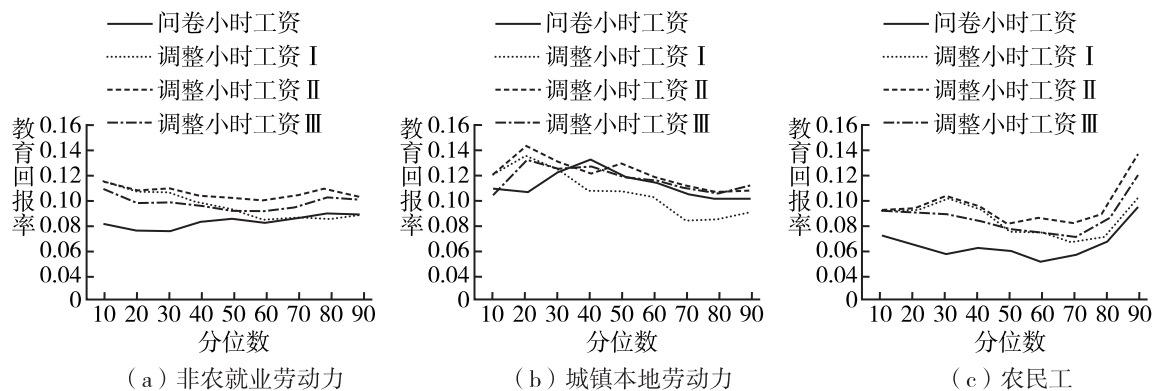


图1 不同社会保险缴费情境下小时工资的教育回报率

注：图中小时工资的教育回报率均根据UQR回归结果进行绘制，且UQR估计得到的教育回报率均具有统计显著性。

其一，由图1(a)可知，问卷小时工资和经社会保险缴费调整后的小时工资在教育回报率上存在明显差异，后者整体高于前者。根据问卷小时工资的回归结果，非农就业劳动力的教育回报率总体上随着收入水平的提高而上升，呈现高收入群体教育回报率更高的特征。而在对问卷小时工资进行社会保险缴费调整后，教育回报率呈现不同的变化趋势。具体而言，按照社会保险缴费基数下限测算的调整小时工资I在教育回报率上表现为，低工资群体的教育回报率较高，随着工资水平的提高，教育回报率整体呈下降趋势。相较之下，按照实际收入进行测算的调整小时工资II和按照应税收入进行测算的调整小时工资III，其教育回报率变化趋势与调整小时工资I存在一定差

异。调整小时工资Ⅱ和调整小时工资Ⅲ的教育回报率并未随着工资水平的提高而整体下降，而是呈现先下降后上升的态势，并在60分位之后出现回升。同时，高工资群体经社会保险缴费调整后的教育回报率，也高于未经调整的问卷小时工资教育回报水平。

这一结果表明，如果不考虑社会保险缴费因素，对低工资群体教育回报率的测算会存在明显低估，即便仅考虑企业缴纳的这部分社会保险，也会显著提升低工资群体的教育回报率。这充分体现出，是否缴纳社会保险对劳动力的当期收入具有重要影响。此外，经社会保险缴费调整后的调整小时工资Ⅱ和调整小时工资Ⅲ，其教育回报率整体呈现一定的U形（“微笑曲线”）特征。在中国非农就业中，低工资群体和高工资群体的教育回报率均高于中等工资水平群体。这说明，中国非农就业中的高工资群体获得了一定的高技能溢价，同时，低工资群体可能因劳动力供给结构和就业相关社会兜底制度获得了结构性溢价或政策性溢价。

其二，通过对比图1（b）和图1（c）可以发现，城镇本地劳动力和农民工的教育回报率存在明显差异。在工资80分位以下区间，城镇本地劳动力的教育回报率普遍高于农民工，而在高分位区间，二者差异较小。同时，二者的教育回报率随工资水平提高的变化趋势也存在显著差异。城镇本地劳动力的教育回报率整体呈倒U形，低工资群体的教育回报率普遍高于高工资群体。经社会保险缴费调整后，以城镇本地劳动力调整小时工资Ⅰ的回归结果为例，教育回报率在低工资分位上的提升幅度更为明显，使城镇本地劳动力教育回报率呈现随工资水平提高而整体下降的趋势。主要原因在于，社会保险缴费基数下限占低工资群体收入的比例较高，社会保险对低工资群体的影响更大；对于缴纳社会保险的低工资群体而言，其初次分配收入会因此得到较大幅度提高。相较之下，社会保险缴费基数下限占高工资群体收入的比例较低，统一按照缴费基数下限调整后，高工资群体调整小时工资Ⅰ的相对差距反而缩小，教育回报率低于未调整时的水平。由此可见，城镇本地劳动力中低工资群体在一定程度上获得了教育溢价，这部分溢价可能源于城镇面向本地劳动力更为完善的兜底保障制度，属于政策层面形成的教育溢价。与城镇本地劳动力不同，由图1（c）可知，农民工的教育回报率呈现较为明显的U形（“微笑曲线”）特征：问卷小时工资的教育回报率在60分位降至最低点，此后逐步上升。经社会保险缴费调整后，农民工教育回报率整体有所上升，在高工资分位上与城镇本地劳动力的教育回报率差距进一步缩小，且整体仍保持“微笑曲线”形态，尽管在不同调整情境下，教育回报率的最低点所处的分位不同。此外，教育回报率在农民工群体内部也存在明显分化：一方面，农民工短缺问题较为突出，为低工资农民工带来教育溢价，使其教育回报率维持在较高水平；另一方面，中等收入农民工尽管受教育年限较高，但其从事的岗位与受教育水平较低的农民工高度趋同，在一定程度上压低了该群体的教育回报率。同时，伴随新旧动能转换，

高技能、高人力资本的农民工的工资增速更快，其教育回报率得到了大幅提高。

## (二) 不同社会保险缴费情境下工资歧视的差异分析

尽管劳动者是否参加社会保险及其社会保险缴费能力主要取决于月工资水平，但从劳动力市场公平性的视角来看，小时工资更能反映劳动力市场的公平程度。因此，本部分重点分析农民工与城镇本地劳动力小时工资差距的 FFL 分解结果（见表 5）。

表 5 农民工与城镇本地劳动力小时工资差距的 FFL 分解结果

变量	10分位	20分位	30分位	40分位	50分位	60分位	70分位	80分位	90分位
问卷小时工资									
总差距	0.228*** (0.076)	0.219*** (0.082)	0.265*** (0.071)	0.269*** (0.067)	0.346*** (0.059)	0.336*** (0.055)	0.339*** (0.036)	0.356*** (0.042)	0.385*** (0.084)
特征效应	0.228*** (0.077)	0.285*** (0.068)	0.325*** (0.076)	0.383*** (0.073)	0.478*** (0.063)	0.598*** (0.092)	0.460*** (0.136)	0.397*** (0.110)	0.542*** (0.106)
结构效应	-0.000 (0.051)	-0.066 (0.046)	-0.060 (0.069)	-0.114 (0.088)	-0.131 (0.084)	-0.262** (0.110)	-0.121 (0.145)	-0.042 (0.112)	-0.157** (0.069)
调整小时工资 I									
总差距	0.468*** (0.070)	0.438*** (0.082)	0.440*** (0.079)	0.410*** (0.058)	0.402*** (0.046)	0.401*** (0.032)	0.391*** (0.029)	0.399*** (0.040)	0.436*** (0.082)
特征效应	0.300*** (0.073)	0.353*** (0.081)	0.428*** (0.080)	0.475*** (0.070)	0.589*** (0.117)	0.509*** (0.141)	0.417*** (0.110)	0.435*** (0.089)	0.483*** (0.106)
结构效应	0.168*** (0.041)	0.085 (0.084)	0.012 (0.095)	-0.064 (0.086)	-0.188 (0.131)	-0.107 (0.154)	-0.027 (0.115)	-0.036 (0.091)	-0.047 (0.094)
调整小时工资 II									
总差距	0.469*** (0.071)	0.443*** (0.083)	0.462*** (0.084)	0.445*** (0.063)	0.464*** (0.056)	0.464*** (0.043)	0.476*** (0.030)	0.515*** (0.050)	0.473*** (0.094)
特征效应	0.302*** (0.071)	0.345*** (0.077)	0.445*** (0.089)	0.501*** (0.081)	0.650*** (0.122)	0.550*** (0.143)	0.538*** (0.114)	0.548*** (0.101)	0.528*** (0.122)
结构效应	0.167*** (0.039)	0.098 (0.076)	0.017 (0.099)	-0.056 (0.092)	-0.186 (0.136)	-0.087 (0.153)	-0.062 (0.114)	-0.033 (0.102)	-0.055 (0.108)
调整小时工资 III									
总差距	0.421*** (0.065)	0.362*** (0.083)	0.405*** (0.078)	0.383*** (0.060)	0.437*** (0.052)	0.427*** (0.042)	0.449*** (0.032)	0.467*** (0.044)	0.464*** (0.092)
特征效应	0.278*** (0.065)	0.316*** (0.077)	0.433*** (0.080)	0.439*** (0.077)	0.629*** (0.110)	0.539*** (0.138)	0.506*** (0.117)	0.503*** (0.100)	0.494*** (0.127)
结构效应	0.143*** (0.034)	0.046 (0.080)	-0.029 (0.088)	-0.056 (0.087)	-0.192 (0.119)	-0.112 (0.150)	-0.056 (0.122)	-0.036 (0.100)	-0.030 (0.107)

注：①括号内为聚类至行业层面的稳健标准误；②\*\*\*和\*\*分别表示 1% 和 5% 的显著性水平。

由表5可知，根据问卷小时工资的回归结果，随着工资水平的提高，农民工与城镇本地劳动力之间的小时工资差距呈扩大趋势，但二者之间的小时工资差距主要源于特征效应，人力资本等个体特征是造成工资差距的核心因素。而且，在各个分位点上，特征效应所解释的工资差距均大于总差距。这意味着，农民工在问卷小时工资层面表现出一定的反歧视特征，即其实际获得的小时工资高于其个体特征所对应的工资水平。

如前文所述，问卷调查中受访者填报的工资多为扣除社会保险缴费后的实发工资，因此，需要根据社会保险缴费情况对其进行调整。经社会保险缴费调整后的工资，增强了参加与未参加城镇职工社会保险的劳动者之间工资的可比性。分解结果表明，以调整小时工资Ⅰ衡量，农民工与城镇本地劳动力之间的工资差距大幅扩大。在10分位上，农民工与城镇本地劳动力之间问卷小时工资（对数值）差距为0.228，而二者调整小时工资Ⅰ（对数值）差距达到0.468。值得注意的是，在10~70分位区间内，二者调整小时工资Ⅰ（对数值）差距整体呈缩小趋势，从0.468下降至0.391。主要原因在于，城镇本地劳动力参加城镇职工基本养老保险的比例达到79.52%，远高于农民工37.05%的参保率，且城镇本地劳动力中低工资群体与农民工中低工资群体之间的参保率差异更为明显。此外，当以实际收入和应税收入作为缴费基数进行工资调整时，调整小时工资Ⅱ和调整小时工资Ⅲ中，农民工与城镇本地劳动力之间的工资差距总体上随着工资水平提高而呈扩大趋势。这是因为，在按照实际收入和应税收入而非社会保险缴费基数下限缴费的情境下，劳动者的缴费额会随着收入水平上升而增加，倒推其得到的初次分配收入也相应提高，从而导致农民工与城镇本地劳动力之间的工资差距扩大。上述结果充分表明，社会保险是劳动者工资性收入的重要组成部分。若忽视社会保险部分而仅分析当期实发工资，不仅可能导致结论偏误，还会忽略劳动者在当期收入与退休收入之间的权衡及选择。

从经社会保险缴费调整后农民工与城镇本地劳动力工资差距的来源来看，尽管调整后二者之间的工资差距有所扩大，但差距仍主要源于特征效应。若以结构效应占总差距的比例衡量歧视程度，则可以发现，相较于农民工在问卷小时工资层面表现出的反歧视特征，经社会保险缴费调整后的工资差距中反歧视程度有所下降（如图2所示）。以调整小时工资Ⅰ为例，农民工在低工资分位上呈现一定的正向结构效应，即仍面临一定程度的歧视：在10分位和20分位上，结构效应占总差距的比例分别为35.90%和19.41%；在30分位上，该比例已接近于0；而在40分位之后，农民工获得了一定程度的反歧视。总体来看，经社会保险缴费调整后，农民工获得的反歧视有所下降，其面临的歧视程度则有所上升。这表明，在未考虑社会保险缴费的情

况下对工资差距进行分析，可能会低估农民工所遭受的歧视，同时高估其获得的反歧视。

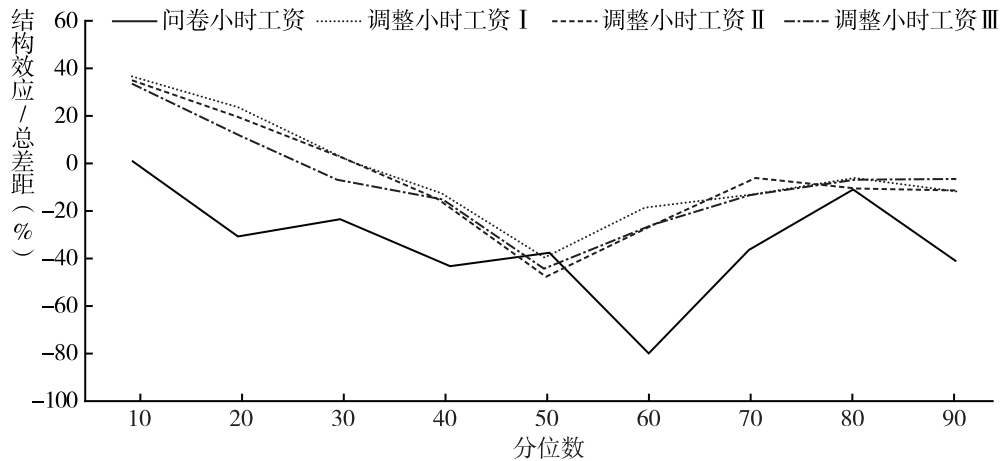


图2 农民工与城镇本地劳动力工资差距的歧视程度分布

农民工与城镇本地劳动力工资差距的分解结果表明，农民工工资水平主要取决于其人力资本等个体特征，农民工在劳动力市场中遭遇歧视而导致工资被压低的现象并不突出。尤其是在依据社会保险缴费情况对工资进行调整后，这一结论仍然成立。这意味着，除低工资分位上的农民工群体外，农民工实际获得的工资性收入在一定程度上已然包含本应缴纳的社会保险部分。结合农民工较低的城镇职工社会保险参保率来看，农民工未参加城镇职工社会保险，可能是其在现行社会保险制度下自主选择的结果。

进一步将FFL分解结果中农民工与城镇本地劳动力工资差距的纯特征效应分解至各解释变量（如图3所示）<sup>①</sup>。由图3可知，受教育年限是特征效应的主要来源，其次是工作经验。值得注意的是，随着收入水平的提高，工作经验的作用逐渐减弱。这表明，在低收入水平上，工作经验是提高工资的重要途径，丰富的工作经验对提高收入水平具有重要作用；而在高收入水平上，工作经验的影响则明显减弱，不仅远低于受教育年限的作用，甚至在高工资分位上呈现负向效应。此外，需要注意的是，在问卷小时工资的10分位上，工作经验对农民工与城镇本地劳动力之间的工资差距几乎不具有影响，但在经社会保险缴费调整后的工资10分位上，工作经验的作用显著提升。这也进一步表明，若忽视社会保险缴费因素，所得结论可能存在较大偏差。

<sup>①</sup> 为突出研究重点并保持图形简洁，图3仅列示受教育年限、工作经验及工作经验的平方项的特征效应贡献。性别、婚姻状况和健康状况等解释变量因贡献度偏低未予展示，并不影响本文的核心结论。

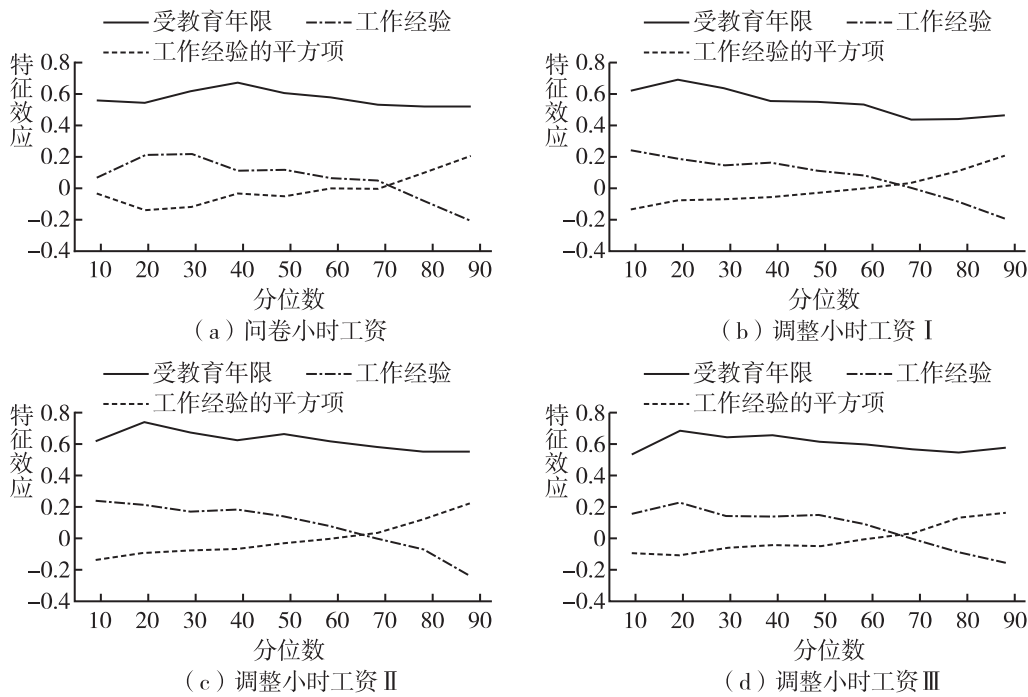


图3 农民工与城镇本地劳动力工资差距的纯特征效应分解结果

## 五、年龄异质性分析与养老保险参保行为拓展分析

### (一) 年龄异质性分析

中国社会保险制度起步较晚，2010年颁布的《中华人民共和国社会保险法》首次在法律层面明确了农民工与城镇职工具有同等参保资格。因此，中年农民工对城镇职工社会保险的了解相对滞后。而随着社会保险制度的推广和完善，不同年龄阶段的农民工对相关制度的了解和认知存在一定的差异。鉴于此，有必要分别探讨不同年龄阶段的农民工与城镇本地劳动力之间工资差距的来源。

青年与中年的年龄划分，尚未形成统一的标准。在由世界银行、世界卫生组织等联合开展的“全球疾病负担”研究中，年龄被划分为15~29岁、30~44岁、45~59岁、60~69岁和70岁及以上，研究显示，酒精对15~44岁和45岁及以上群体身体健康的影响存在显著差异（Murray and Lopez, 1996）。根据中国法定劳动年龄的规定，年满16周岁被视为劳动年龄人口。此外，在人口老龄化定义中，老年人口的划分标准通常为年龄在60岁及以上或65岁及以上。综合上述标准，并结合2023年CSS受访者的年龄分布，本文将青年群体界定为年龄在18~44岁，将中年群体界定为年龄在45~64岁。由于调整小时工资Ⅱ、调整小时工资Ⅲ与调整小时工资Ⅰ的分析结果具有较强的一致性，在考察不同年龄群体中农民工与城镇本地劳动力工资差距的来源时，

除问卷小时工资外，仅报告调整小时工资 I 的 FFL 分解结果<sup>①</sup>。

其一，对青年群体中农民工与城镇本地劳动力工资差距的来源展开分析。分解结果显示，青年农民工与青年城镇本地劳动力的问卷小时工资差距普遍呈现反歧视效应，即由个体特征差异带来的特征效应大于总差距。经社会保险缴费调整后，在调整小时工资 I 的 70 分位以下，青年农民工与青年城镇本地劳动力之间的工资差距明显扩大，但在 70 分位以上，二者之间的工资差距并未发生明显变化。这反映出，青年农民工中的中低收入群体参加城镇职工社会保险的比例低于青年城镇本地劳动力，但高收入群体的参保比例与青年城镇本地劳动力相近。此外，在调整小时工资 I 的 20 分位以上，青年农民工获得了反歧视（特征效应大于总差距），但在 10~20 分位上，青年农民工仍面临一定程度的工资歧视。

其二，对中年群体中农民工与城镇本地劳动力工资差距的来源展开分析。由分解结果可知，中年农民工与中年城镇本地劳动力的工资差距高于青年群体。在问卷小时工资的 40 分位以下，中年农民工与中年城镇本地劳动力工资差距中的结构效应占比超过 10%，其中，10 分位上的结构效应占比接近 60%，表明工资歧视问题仍较为突出。而在 60 分位以上，二者之间的工资差距则主要源于特征效应，甚至呈现一定的反歧视特征。值得注意的是，经社会保险缴费调整后，中年农民工与中年城镇本地劳动力之间的工资差距显著扩大，且低工资分位上工资差距扩大幅度更为明显，而在 80 分位和 90 分位上，工资差距扩大幅度相对偏小。这一变化趋势与青年群体在社会保险缴费调整后的工资变化趋势基本一致。上述结果同样表明，中年农民工中的中低收入群体参加城镇职工社会保险的比例较低，而高收入群体的参保率较高，故高收入群体受社会保险缴费调整的影响较小。从工资差距的来源来看，中年农民工普遍面临工资歧视，同时，经社会保险缴费调整后所体现的工资歧视程度，明显高于问卷小时工资所反映的水平。

分年龄阶段考察农民工与城镇本地劳动力的工资差距来源后可以发现，相较于青年农民工而言，中年农民工面临的工资歧视更为普遍且程度较高。青年群体中农民工与城镇本地劳动力之间的工资差距主要来自特征效应，而中年农民工则面临工资歧视。《2024 年农民工监测调查报告》显示，2024 年中国农民工的平均年龄为 43.2 岁，40 岁以上群体占比高达 56.3%<sup>②</sup>。而且，近年来农民工的平均年龄持续提高。因此，中年农民工所面临的工资歧视问题应得到进一步关注。

① 限于篇幅，青年和中年群体中农民工与城镇本地劳动力小时工资差距的 FFL 分解结果，详见《中国农村观察》网站或中国知网本文附录中的附表 1。

② 资料来源：《2024 年农民工监测调查报告》，[https://www.stats.gov.cn/sj/zxfb/202504/t20250430\\_1959523.html](https://www.stats.gov.cn/sj/zxfb/202504/t20250430_1959523.html)。

## （二）养老保险参保行为拓展分析

根据前文对农民工与城镇本地劳动力工资差距及其来源的分析可以发现，经社会保险缴费调整后，中低收入群体中农民工与城镇本地劳动力之间工资差距的扩大幅度更为明显，而个体特征仍然是二者工资差距的主要来源。这表明，农民工的个体特征在劳动力市场中得到了较为充分的体现，其与城镇本地劳动力之间的工资差距主要源于可解释的特征效应。由此可以推断，农民工的低参保率更可能是其在现行社会保险制度下自主选择的结果，即其贴现因子较高，故在当期收入与养老金收入之间更倾向于选择当期收入。鉴于此，本文从家庭支出需求的视角出发，考察若家庭预期有较大消费支出会如何影响劳动者的养老保险参保行为。一般而言，当家庭预期消费支出较高时，相较于未来的养老保障，家庭往往更加重视当期收入；同时，对于工资水平较低的劳动者而言，其对当期收入的需求更为迫切，因此，更倾向于不参保，或选择缴费负担较低的城乡居民基本养老保险。本部分采用多元Logit模型，重点考察家庭中是否有未婚男性、是否有在校学生以及家庭医疗支出负担是否较重对养老保险参保行为的影响，并将调整月工资 $I$ 、受教育年限、性别和年龄作为解释变量，同时控制地区变量和行业变量<sup>①</sup>。其中，调整月工资 $I$ 用于控制劳动者的参保能力。养老保险参保行为作为被解释变量，包含三类情况：不参加养老保险、参加城镇职工基本养老保险和参加城乡居民基本养老保险。

回归结果显示，家庭中有未婚男性会显著降低劳动者参加城镇职工基本养老保险的概率<sup>②</sup>。以调整月工资 $I$ 的中位数为界，进一步将样本划分为中高收入组和中低收入组进行回归可以发现，家庭中有未婚男性对劳动者参保行为的影响，主要体现在农民工中高收入组。当家庭中有未婚男性时，中高收入的农民工参加城镇职工基本养老保险的概率显著降低，不参加养老保险的概率显著上升。由此可见，家庭预期面临较高婚姻支出，是影响农民工群体，尤其是中高收入的农民工群体养老保险参保行为的重要因素。在中低收入组中，家庭中是否有未婚男性并不会对劳动者参保行为产生显著影响。

当家庭中有在校学生时，劳动者不参加养老保险的概率显著降低<sup>③</sup>。需要注意的是，该效应在中高收入群体中呈现明显的群体异质性。具体而言，在中高收入组中，

① 在进一步分析中，由于需要按照收入水平进行分组回归，基于对组内样本量充分性的考量，本文仅控制华北、东北、华东、中南、西南、西北六大区域和第二产业，分别作为地区层面和行业层面的控制变量。

② 限于篇幅，家庭中有未婚男性对养老保险参保行为影响的回归结果详见《中国农村观察》网站或中国知网本文附录中的附表2。

③ 限于篇幅，家庭中有在校学生对养老保险参保行为影响的回归结果详见《中国农村观察》网站或中国知网本文附录中的附表3。

家庭中有在校学生并不会对城镇本地劳动力的参保行为产生显著影响，却会显著提高农民工参加城镇职工基本养老保险的概率，并显著降低其不参加养老保险的概率。这可能与农民工随迁子女的就学准入门槛有关，一些地区将连续缴纳一定期限（如一年及以上）的社会保险证明作为随迁子女入学条件之一。因此，为保障子女能够随迁至农民工就业地顺利就学，农民工更倾向于参加城镇职工基本养老保险。

当受访者认为家庭医疗负担较重时，城镇本地劳动力和农民工的参保行为呈现显著差异<sup>①</sup>。具体来看，家庭医疗负担较重会显著提高城镇本地劳动力参加城镇职工基本养老保险的概率，却会显著降低农民工参加城镇职工基本养老保险的概率。这可能与城镇职工基本养老保险和基本医疗保险共同缴纳的特征有关。对于城镇本地劳动力而言，较重的医疗负担会强化其通过参保分担医疗支出压力的动机，从而提高其参加城镇职工基本养老保险的概率。然而，对于农民工而言，较重的医疗负担可能削弱其缴费能力，从而降低其参加城镇职工基本养老保险的概率，并提高其不参加养老保险的可能性。进一步按收入水平分组后可以发现，在中高收入组中，较重的医疗负担并不会显著影响城镇本地劳动力的参保行为，却会显著提高农民工参加城乡居民基本养老保险的概率，并降低其参加城镇职工基本养老保险的概率。在中低收入组中，较重的医疗负担会显著提高城镇本地劳动力参加城镇职工基本养老保险的概率，且会显著提高农民工不参加养老保险的概率。

此外，调整月工资 I 对劳动者参加养老保险具有显著影响。回归结果表明，当调整月工资 I 提高时，农民工和城镇本地劳动力参加城镇职工基本养老保险的概率均显著提高，不参加养老保险和参加城乡居民基本养老保险的概率均显著下降。尤其在中低收入组中，调整月工资 I 对农民工和城镇本地劳动力参保行为的影响更为明显。由此可见，提高工资水平是促进低工资劳动者参加城镇职工基本养老保险的重要途径。

## 六、结论与启示

工资性收入是中国居民的主要收入来源，工资分配的公平性既是按贡献决定报酬这一初次分配原则的重要体现，也是构建公平合理收入分配格局的重要基础。从生命周期视角来看，养老金收入是在职收入公平性的进一步延续。现有基于微观调查数据对劳动力市场工资歧视的研究，大多直接采用受访者填报的工资数据进行分析。然而，受访者在接受问卷调查时，基于理解和填写习惯，其填报的工资多为扣除个人所

<sup>①</sup> 限于篇幅，家庭医疗支出负担较重对养老保险参保行为影响的回归结果详见《中国农村观察》网站或中国知网本文附录中的附表4。

得税与社会保险缴费后的实发工资。由此，如果直接使用问卷工资开展研究，实质上是将参加城镇职工社会保险群体的税后实发工资，与参加城乡居民社会保险或未参加社会保险群体的初次分配收入进行比较，但显然二者并不具有可比性，由此得到的实证结论可能出现明显偏误。若政策建议建立在这种缺乏可比性，乃至存在偏误的微观分析结果之上，便难以具备应有的针对性，甚至可能导致政策着力点发生偏离。促进高质量充分就业和完善收入分配制度，是中国“十五五”时期扎实推进全体人民共同富裕的重要目标。而这两项目标都与劳动力市场特征及劳动者的社会保险参保行为密切相关。鉴于此，本文采用2023年中国社会状况综合调查数据，依据各省份社会保险缴费政策，结合不同测算情境对微观调查的工资性收入进行调整，并采用UQR和FFL分解方法，考察农民工与城镇本地劳动力之间问卷工资差距和经社会保险缴费调整后的工资差距及其来源。

本文研究发现：首先，从更能体现劳动力市场效率的小时工资来看，经社会保险缴费调整后，农民工与城镇本地劳动力小时工资的教育回报率的差距有所缩小。分别来看，调整后城镇本地劳动力的教育回报率整体呈倒U形分布，低工资群体的教育回报率普遍高于高工资群体；而调整后的农民工教育回报率有所上升，整体保持“微笑曲线”形态，其教育回报率在高工资分位上与城镇本地劳动力的教育回报率差距进一步缩小。其次，农民工与城镇本地劳动力问卷小时工资差距来源中特征效应占主导地位，农民工在问卷小时工资层面表现出一定的反歧视特征。经社会保险缴费调整后，二者之间的工资差距虽有所扩大，但仍以特征效应为主要来源，同时，农民工获得的反歧视有所下降，工资歧视在一定程度上依然存在。这表明，在未考虑社会保险缴费的情况下对工资差距进行分析，可能会低估农民工所遭受的歧视，同时高估其获得的反歧视。再次，受教育水平是工资差距中可解释的特征效应的主要来源，工作经验次之。然而，随着收入水平的提高，工作经验的作用逐渐减弱。分年龄阶段来看，中年农民工面临的工资歧视明显高于青年农民工。青年群体中农民工与城镇本地劳动力之间的工资差距主要来自特征效应，而中年农民工则面临工资歧视，且在低工资分位区间面临的歧视程度更高。最后，家庭预期支出和劳动者工资水平对养老保险参保行为具有显著影响。当家庭中有未婚男性（预期有较高婚姻支出）时，中高收入组中农民工参加城镇职工基本养老保险的概率显著降低，而不参加养老保险的概率则显著上升；当家庭中有在校学生（预期有较高教育支出）时，中高收入组中农民工参加城镇职工基本养老保险的概率显著提高，不参加养老保险的概率则显著下降；较重的医疗负担会显著提高城镇本地劳动力参加城镇职工基本养老保险的概率，同时，会显著降低农民工参加城镇职工基本养老保险的概率；此外，当收入水平提高时，特别是当工资较低群体的收入水平提高时，劳动者参加城镇职工基本养老保险的概率会显著提高。

基于上述研究结论，本文得到以下政策启示：

第一，健全与常住人口相适应的教育资源配置机制，以“投资于人”赋能人的全面发展。一方面，应完善面向农民工的继续教育和职业技能培训体系，着力提升其人力资本水平和劳动力市场竞争力，特别是加大对中年农民工的培训支持力度，推动终身职业技能培训制度不断健全；另一方面，要切实提高农民工子女的就学保障水平，促进教育机会公平。从工资差距的分解结果来看，受教育水平是农民工与城镇本地劳动力工资差距的重要来源。这也意味着，促进教育公平既是缩小两类群体之间收入差距的重要方式，也是缓解代际收入不平等传递的重要路径。因此，要消除或缓解两类群体之间的教育差距，亟须进一步完善与常住人口相适应的教育资源配置机制。这不仅是缩小劳动力市场不平等的重要举措，也是加快推动中国人口红利向人才红利转化的关键所在。

第二，深化完善社会保障体系，优化社会保险制度设计。立足现行城镇职工社会保险、灵活就业人员社会保险与城乡居民社会保险制度框架，加快完善多层次社会保险制度，持续健全社会保险关系转移接续机制。在增强社会保障整体包容性的同时，切实降低农民工参加城镇职工社会保险的参保负担，稳步提高其参保意愿和参保水平，推动形成覆盖范围更广、制度衔接更顺畅、适配性更强的社会保障体系。

第三，加强劳动者就业权益保障，营造公平有序的劳动力市场环境。伴随中国新旧动能转换的持续推进和新质生产力的加速发展，劳动力市场结构性矛盾日益凸显，劳动者在就业机会获得、工资分配与劳动保障等领域的不平等现象依然存在。本文研究发现，中年农民工群体面临的工资歧视问题仍较为突出。因此，应进一步完善城乡劳动者平等就业制度，加大劳动监察执法力度，依法纠正各类就业歧视，切实保障农民工劳动报酬权益，健全有利于更充分更高质量就业的促进机制，不断增强广大劳动者的获得感幸福感安全感。

#### 参考文献

蔡昉，2010：《中国发展的挑战与路径：大国经济的刘易斯转折》，《广东商学院学报》第1期，第4-12页。

陈昊、赵春明、杨立强，2017：《户籍所在地“反向歧视之谜”：基于收入补偿的一个解释》，《世界经济》第5期，第173-192页。

邓曲恒，2007：《城镇居民与流动人口的收入差异——基于Oaxaca-Blinder和Quantile方法的分解》，《中国人口科学》第2期，第8-16页。

洪甘霖、赵宗胤、钱文荣，2023：《社会养老保险对农地转出影响的再审视——基于心理账户理论视角的分析》，《中国农村经济》第12期，第65-84页。

贾晗睿、詹鹏、李实，2021：《“多轨制”养老金体系的收入差距——基于中国家庭收入调查数据的发

现》，《财政研究》第3期，第101-114页。

江求川、鲁元平，2024：《户籍歧视导致的收入差距依然存在吗——基于机器学习方法的再讨论》，《中国农村经济》第9期，第101-120页。

李培林、李炜，2007：《农民工在中国转型中的经济地位和社会态度》，《社会学研究》第3期，第1-17页。

李实、赵人伟、高霞，2013：《中国离退休人员收入分配中的横向与纵向失衡分析》，《金融研究》第2期，第1-18页。

王春光，2006：《农村流动人口的“半城市化”问题研究》，《社会学研究》第5期，第107-122页。

王美艳，2005：《城市劳动力市场上的就业机会与工资差异——外来劳动力就业与报酬研究》，《中国社会科学》第5期，第36-46页。

魏后凯、苏红键，2013：《中国农业转移人口市民化进程研究》，《中国人口科学》第5期，第21-29页。

邢春冰、屈小博、杨鹏，2021：《农民工与城镇职工工资差距演变及原因分析》，《经济学动态》第5期，第64-78页。

徐志刚、宁可、钟甫宁、纪月清，2018：《新农保与农地转出：制度性养老能替代土地养老吗？——基于家庭人口结构和流动性约束的视角》，《管理世界》第5期，第86-97页。

余向华、陈雪娟，2012：《中国劳动力市场的户籍分割效应及其变迁——工资差异与机会差异双重视角下的实证研究》，《经济研究》第12期，第97-110页。

岳希明、徐静，2024：《收入再分配政策调节机制的探讨》，《经济研究》第4期，第173-190页。

张培文、耿献辉、卢华，2023：《基本公共服务何以影响县域农民工市民化》，《财政研究》第12期，第34-48页。

章莉、吴彬彬，2019：《就业户籍歧视的变化及其对收入差距的影响：2002—2013年》，《劳动经济研究》第3期，第84-99页。

朱玲，2014：《转向适应市场经济运行的社保体系》，《劳动经济研究》第4期，第3-20页。

Ando, A., and F. Modigliani, 1963, "The 'Life Cycle' Hypothesis of Saving: Aggregate Implications and Tests", *The American Economic Review*, 53(1): 55-84.

Firpo, S., N. Fortin, and T. Lemieux, 2009, "Unconditional Quantile Regressions", *Econometrica*, 77(3): 953-973.

Firpo, S., N. Fortin, and T. Lemieux, 2018, "Decomposing Wage Distributions Using Recentered Influence Function Regressions", *Econometrics*, 6(2), <https://doi.org/10.3390/econometrics6020028>.

Murray, C., and A. Lopez, 1996, *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020*, Cambridge, MA: Harvard University Press, 28-30.

Rios-Avila, F., 2020, "Recentered Influence Functions (RIFs) in Stata: RIF Regression and RIF Decomposition", *The Stata Journal*, 20(1): 51-94.

Zhang, D., 2020, "The Evolution of the Wage Gap Between Rural Migrants and the Urban Labour Force in Chinese Cities", *Australian Journal of Agricultural and Resource Economics*, 64(1): 55-81.

Zhu, J., S. Jin, Z. Tang, and T. Awokuse, 2022, "The Effect of Pension Income on Land Transfers: Evidence from Rural China", *Economic Development and Cultural Change*, 71(1): 333-371.

## Rethinking the Wage Gap Between Rural Migrant Workers and Urban Local Workers: An Empirical Analysis from the Perspective of Social Insurance Participation

SUN Jingfang

(Institute of Economics, Chinese Academy of Social Sciences)

**Summary:** With the ongoing transformation of labor market structures, existing studies suggest that the discrimination faced by rural migrant workers has weakened or even tended to disappear. However, relatively few studies on wage discrimination against rural migrant workers incorporate social insurance participation. As an important component of labor compensation, social insurance contributions are often subject to notable statistical underreporting in micro-level wage data. Such underreporting can introduce bias into analyses of the wage gap between rural migrant workers and urban local workers. Specifically, employee social insurance is withheld and remitted by employers on behalf of employees, whereas resident social insurance is paid directly by individuals. In this context, workers participating in employee social insurance typically report post-contribution disposable income in micro surveys, while those participating in resident social insurance report wages closer to pre-distribution income. If social insurance contributions are excluded from reported wages, the data become non-comparable between employee and resident social insurance participants.

This study utilizes data from the 2023 Chinese Social Survey and based on provincial social insurance contribution policies, adjusts reported wages under different social insurance contribution scenarios to recover labor income more consistent with primary distribution. On this basis, the study examines the wage gap between rural migrant workers and urban local workers and its sources. The results indicate that the wage gap between the two groups is primarily driven by differences in individual characteristics, among which educational attainment constitutes the core source, followed by work experience. Using hourly wages adjusted for social insurance contributions and aligned with primary distribution, the wage gap remains mainly driven by the characteristics effect; however, reduced anti-discrimination for rural migrant workers indicates that wage discrimination persists to some extent. In terms of age heterogeneity, the wage gap between rural migrant workers and urban local workers among younger workers is mainly attributable to the characteristics effect, whereas middle-aged rural migrant workers face wage discrimination.

This study yields the following implications: under current labor market conditions, rural migrant workers do not experience significant wage discrimination, even when primary distribution of income—back-calculated by adjusting for social insurance contributions—is used as the benchmark. The relatively low participation rate of employee social insurance among rural migrant workers is more likely to reflect self-selection under the existing social insurance system. In addition, greater attention should be directed toward wage discrimination faced by middle-aged rural migrant workers and issues related to pension insurance.

**Keywords:** Social Insurance; Primary Distribution of Income; Wage Gap; Rural Migrant Workers

**JEL Classification:** J31; J71

(责任编辑:柯 宓)