

流动人口社会身份认同、过度劳动与城乡差异*

孙文凯 王格非

摘要：我国流动人口过度劳动问题非常严重，而现有相关研究还相对缺乏。本文基于2017年中国国家卫生与健康委员会流动人口调查监测数据，分析流动人口城市身份认同对过度劳动的影响。我们认为如果流动人口融入城市，将可能由于示范效应等途径而降低其过度劳动可能。在基础OLS分析之后，本文使用了工具变量和差分法进一步识别二者间因果关系，发现流动人口产生城市本地身份认同能够有效缓解过度劳动发生，并减少过度劳动时间，而城乡户籍流动人口间差异不明显。向上模仿、未来去向和本地适应都是产生影响的机制。本文还根据流动人口的特征对过度劳动情况进行了异质性分析。除对过度劳动情况的广度进行分析外，本文对过度劳动时间即过度劳动的深度也进行了研究，并得到了一致性的结论。本文的政策含义是加强流动人口社会融合可以减少过度劳动并改善流动人口健康。

关键词：过度劳动；流动人口；社会身份认同；工具变量；城乡差异

中图分类号：F245 JEL：J15 J61 J83

Social Identity, Overwork and Urban-rural Difference of Migrant Workers

SUN Wenkai WANG Gefei

(Renmin University of China, Beijing, China)

Abstract: The problem of overwork of migrant workers is very serious in China. This phenomenon, however, is still less studied. Based on the China Migrants Dynamic Survey of the National Health Commission of China in 2017, this paper analyzed the impact of social identity of migrant worker on overwork. We believe that if the migrants integrated into the city, it may reduce the possibility of overwork

*孙文凯、王格非，中国人民大学经济学院，邮政编码：100872，电子邮箱：sunwk@ruc.edu.cn，wgfsofia@163.com。本研究受国家自然科学基金(71603266)和国家社科基金重大项目(18ZDA081)支持。感谢审稿人的宝贵意见，作者文责自负。

due to demonstration effect. After the basic OLS analysis, we use Instrumental Variable Method and differencing model to identify the causal relationship between them. We find that the local identity of migrant population can significantly alleviate the occurrence of overwork and reduce the overwork time. There is no significant difference between migrant workers with urban and rural hukou. We also found that upward imitation, future uncertainty and local adaptation are all influential mechanisms. The policy implication of this paper is that promoting the social integration of migrant population can reduce overwork and improve their health.

Keywords: Overwork; Migrants; Social Identity; Instrumental Variable; Urban-Rural Difference

一、引言

近年来, 过度劳动现象在我国引起了普遍关注。根据中国人力资源和社会保障部发布的《2016年中国劳动统计年鉴》^①, 2015年城镇就业人员的周平均劳动时间为45.5小时, 其中有30.4%的就业人员每周平均劳动时间在48小时以上。对比同年经合组织(OECD)数据, OECD国家2015年周平均劳动时间仅为36.7小时^②, 中国远超出OECD国家平均值。国际经验已经显示过度劳动会严重影响劳动者的身体健康、精神状态、家庭关系以及主观幸福感(Caruso, 2006; Dahlgren et al, 2006)。十九大报告提出“全面推进健康中国战略”, 这要求对我国劳动人口过度劳动问题额外重视。同时, 我国流动人口的过度劳动现象更为严重。根据国家卫生和健康委员会2017年调查数据, 流动就业人口周平均劳动时间为56.8小时, 58.2%的流动就业人口周平均劳动时间超过50小时, 而2015年这一比例还只有52.2%。缓解流动人口过度劳动情况, 是提升流动人口健康水平、增强流动人口劳动保护的重要任务。

目前国内学术界已有较多针对过度劳动成因的研究, 但这些研究较少将目光集中在问题更严重的流动人口群体上, 也没有研究从身份认同这一心理效应来解释过度劳动现象。社会身份认同(social identity)理论的起源可以追溯到Tajfel & Turner (1979)。身份认同代表人们自我定位所属群体身份, 进而自我行为服从此群体行为规范。在我国, 身份认同理论被广泛应用于流动人口能否融入当地社会的问题分析上(杨菊华, 2010)。由于流动人口在城市劳动和生活却一直受到区别对待, 很多研究都指出了农民工在城市身份认同存在困境(张静, 2010; 崔岩, 2012)。流动人口往往只认为自己是城市的“过客”, 在城市社会身份认同的差异会导致经济行为的差异(Chen & Li, 2009)。一个引申的猜想是: 正是由于流动人口这种“过客”的心态, 促使其短期内过度投入努力, 以获取较高回报为返乡或继续流动准备。对应的问题就是: 流动人口产生城市本地社会身份认同是否能够有效缓解他们的过度劳动情况? 回答这个问题不仅有助于我们了解流动人口的劳动行为, 也有助于促进流动人口健康劳动相关政策的评价设计。

本文使用国家卫生和健康委员会2017年流动人口监测数据定量分析流动人口社会身份认同对过度劳动情况的影响。除了用工具变量方法解决可能的内生性问题, 基于城乡户籍差异, 本文还检

^①中华人民共和国统计局. 中国统计年鉴[M]. 北京: 中国统计出版社, 2016.

^② 数据来自 OECD 数据库, 访问时间 2018. 12. 13, 链接 <https://stats.oecd.org/Index.aspx?DataSetCode=ANHRS>.

验不同户口类型流动人口的社会身份认同对过度劳动情况的影响差异，这有利于进一步识别影响来源。本文对社会身份认同对过度劳动的其他影响机制进行了分析，证明向上流动、未来去向和本地适应影响了流动人口的过度劳动情况。另外本文还根据流动人口各项特征对过度劳动情况的异质性进行了讨论。除对过度劳动广度的讨论外，本文将过度劳动时间作为解释变量对过度劳动深度也进行了讨论。

本文的创新之处在于：第一，本文将流动人口作为研究过度劳动问题的主要研究对象，并区分了城城流动和城乡流动。我国流动人口数量庞大，但存在着就业质量较低，收入不确定性更高的特点。单独对流动人口的过度劳动问题进行研究，有助于促进流动人口劳动保护政策的设计，能够有针对性地提升劳动者的劳动健康；第二，本文从身份认同这一心理效应出发进行研究，从就业质量与劳动健康的角度补充了有关流动人口社会认同困境的研究，有助于进一步了解流动人口的流入地社会融合相关问题，促进流动人口相关劳动政策的评价和设计；第三，本文使用了多种工具变量处理了在处理身份认同过程中可能遇到的内生性问题，讨论了可能由于流动人口各项特征引起的过度劳动异质性，并加入劳动时间作为被解释变量对过度劳动问题的广度进行了研究，增加了结论的可靠性和稳健性。

本文其余部分内容安排如下：第二部分进行相关文献简要回顾，第三部分描述数据并介绍本文所用统计方法；第四部分首先用 OLS 模型进行基准分析，之后对潜在内生性问题进行处理，对社会身份认同对过度劳动的影响机制进行了分析，并对过度劳动情况的异质性进行了讨论，最后以流动人口的过度劳动时间为被解释变量进一步研究过度劳动深度问题；第五部分总结全文。

二、文献综述

大量学者利用不同国家的数据对各国过度劳动的现状以及发展趋势进行了总结，发现过度劳动在世界范围内有普遍增长趋势。Kuhn & Lozano (2008) 利用美国 1976-2006 年近期人口调查数据 (current population survey) 发现美国男性劳动者的工作时间持续延长，且男性劳动者每周工作小时超过 48 小时的比例也在增加，这一现象在高教育水平、高收入和高龄群体中更为严重。根据《中国劳动统计年鉴 2016》的统计数据，中国的劳动时间分布略有不同，其中教育水平较低的中青年劳动者平均周劳动时间最长。Drago et al (2005) 利用澳大利亚数据证明劳动者过度劳动具有持久性。国内学者近年来也开始关注过度劳动问题，并从个人、用人单位和社会环境等多方面讨论了我国过度劳动问题的成因。肖红梅 (2014) 通过调查数据分析发现性别、工作性质等因素对解释个体过度劳动有显著作用。郭凤鸣、曲俊雪 (2016) 利用中国营养与健康调查数据发现我国过度劳动情况呈现加剧态势，并且男性、低技能和私营企业劳动者过度劳动的概率较高。少量研究发现流动人口的过度劳动情况更为严重。Park & Wang (2010) 利用 2005 年中国城市劳动力调查数据发现，城镇居民和流动人口的劳动时间存在着显著差异，大城市流动人口月平均劳动时间为 283 小时，远高于对应大城市城镇居民的 184 小时；小城市流动人口的月平均劳动时间为 250 小时，远高于小城市城镇居民的 183 小时。郭凤鸣、张世伟 (2018) 统计指出，低收入农民工群体存在着严重的过度劳动问题，其中男性农民工群体中每周工作时间大于 50 小时的比例大于 70%。

已有部分文献针对过度劳动影响因素进行了理论或实证研究。Wolfe (1997) 以及 Ellingsen & Johannesson (2007) 认为劳动时间由劳动者的内在激励和外在激励同时决定。一些学者提出，将劳动视为对未来工资的前瞻性投资行为能够解释过度劳动的发生 (Golden, 2008)。Michelacci & Pijoan-Mas (2007) 也认为劳动时间的延长反映了劳动者对未来职业提升的期望，是一种主动行为，收入不平等越大这种行为会越显著。劳动时间的延长也可能是应对未来失业或者降薪风险的措施，职业或收入的不确定性增加会加剧过度劳动 (Bluestone & Rose, 1998)。郭凤鸣、张世伟 (2018) 利用 2012 和 2013 年流动人口动态监测数据结合城市宏观数据，证实最低工资标准的提升增加了女

性农民工群体的失业风险，导致了这部分农民工群体过度劳动时间的延长。李君甫和许多（2018）利用中国国家卫健委 2016 年流动人口动态监测调查数据，发现雇主提供住房的居住方式更容易造成农民工的过度劳动。郭凤鸣、牟林（2019）发现政府培训和提升工资能够缓解流动人口过度劳动。王静、王欣（2013）的统计分析发现家庭压力、就业稳定性、收入水平和未来愿景会影响农民工过度劳动。在我国，流动人口面临着更为严重的职业和收入不确定性（邓曲恒、2007；马小红、段成荣、郭静，2014），并且这一现象主要发生在乡—城流动人口中（杨菊华，2011），这种不确定性可能对流动人口整体的过度劳动有解释力。

已有部分研究与社会身份认同影响劳动时间的视角相关。这个视角下身份认同可能对流动人口产生影响的渠道主要是模仿效应或参照点效应。首先，流动人口可能通过参照点的模仿效应减少过度劳动时间。较多针对社会身份认同的研究证实了流动人口的本地社会身份认同能够影响参照群体的经济行为，如消费、学习和亲社会行为等（Fang & Loury, 2005; Chen & Li, 2009; Afridi et al, 2015; 王湘红、陈坚，2016）。杨菊华（2009）将流动人口的社会融入分为经济整合、文化接纳、行为适应和身份认同四个维度。当流动人口对流入地有很强的认同感时，意味着他们在流入地产生了较强的经济整合和行为适应，具有与流入地城镇居民类似的行为特征。张文宏、雷开春（2009）认为在有较好的社会身份认同时，流动人口在流入地的经济、文化、行为等都会以目的地居民为参照标准，言行举止更接近当地人。由于城镇居民劳动时间相对少，可以预期城市中本地认同感更强的流动人口将通过模仿效应减少工作时间。国外研究也提供了类似证据，比如 Kahanec & Shields（2010）利用德国移民数据证明未来不倾向于留在工作地的移民更有可能发生过度劳动。同时，参照点效应还可以通过社会比较视角得到与模仿机制不同的影响方向预测。已有研究显示劳动者会出于证明个人的成功或提升相对社会地位的目的进行消费，所以为满足更高水平的消费，劳动者会延长他们的劳动时间（Altman, 2001）。当收入不平等情况加剧时，那些收入较低的劳动者可能会出于维持相对消费水平或模仿富裕群体消费行为的目的而延长工作时间（Bowles & Park, 2005）。如果流动人口认为自己是城市本地人，那么他们将更多和城市居民比较并进行高消费，并且由于当前较低收入必须通过过度劳动来支撑其消费。而差距较大的认为自己仍是老家人的流动人口反倒可能减少过度劳动，放弃竞争性比较（王湘红等，2012；孙文凯等，2019）。流动人口城市本地人身份认同对其过度劳动可能提出新解释，但以上两个机制的解释综合影响方向仍需检验。

尽管目前国内已有一些研究表明社会身份认同的差异能够造成流动人口经济行为的差异，但尚未有直接针对流动人口社会身份认同对劳动时间影响的研究。本文目的即是探讨城市本地社会身份认同对流动人口过度劳动影响并讨论影响机制。

三、 数据和分析方法

（一） 数据来源

本研究使用数据来自 2017 年国家卫健委进行的流动人口动态监测，这是目前国内规模最大的专门针对流动人口进行的调查。该数据按照随机原则在全国 31 个省（区、市）和新疆生产建设兵团流动人口集中流入城区抽取样本点，调查对象主要为 15-59 周岁的流入人口，调查内容包括流动人口的个人信息、流动信息和社会融合情况，也包括流动人口就业特征、家庭特征以及其他基本公共卫生服务情况。问卷调查中“本周工作小时数”可以作为计算流动人口过度劳动情况计算的主要依据。排除调查对象中未参与劳动的群体后，本文最终共得到 115384 个分析样本。

由于这一数据中缺乏流入地城镇户籍居民的劳动时间数据，本文还将根据流入地城市将流动人口数据与中国劳动力动态调查（CLDS）2016 年数据进行了匹配，其中有效的匹配数据样本 1948 个，研究中主要使用了匹配样本的户籍信息和劳动时间数据。

（二）模型设定

本文采用的基本分析方法为多元回归分析，控制其他因素后考察流动人口流入地城市社会身份认同对过度劳动情况的影响，基准 OLS 回归方程设定如下：

$$overwork_i = \alpha_i + \alpha_1 SocialIdentity_i + \alpha_0 X_i + \varepsilon_i \quad (1)$$

其中 $overwork_i$ 为流动人口 i 是否存在过度劳动的二值哑变量。^①Jacobs & Green (1998) 在有关美国过度劳动情况的研究中，将每周劳动时间超过 50 小时作为识别过度劳动的标识。在 Cha & Weeden (2014) 针对美国过度劳动与性别差异的研究中，同样直接使用劳动时间大于等于 50 小时作为过度劳动的标志，郭凤鸣、曲俊雪 (2016) 利用中国健康与营养调查数据 (CHNS) 进行的过度劳动研究中，也使用了这一定义。本文在基础分析中沿用了这一定义^②，若流动人口周劳动时间大于等于 50 小时， $overwork_i$ 取值为 1，否则为 0。 $SocialIdentity_i$ 代表了流动人口的本地城市人身份认同。现有研究表明流动人口是否认同自己属于“本地人”这一社会身份可以被作为衡量其在居住地社会融入的重要标志 (张文宏、雷开春，2009；崔岩，2012)。 X_i 是其他可能影响流动人口劳动时间的个人特征、家庭特征和工作特征，个人特征包括性别、年龄、户口性质、教育背景、健康状况等，家庭特征包括家庭子女数量以及家庭土地承包情况等，工作特征包括工作行业和工资率。 ε_i 为随机误差项。

由于调查样本都在城市地区，而流动人口中既有农业户籍人口也有非农户籍人口。孙文凯 (2017) 认为，户籍制度影响了流动人口的消费、就业等行为，并且在收入较高的城市更难实行户籍开放。因此对于农业户籍流动人口，问卷中的“本地人”身份认知包含两重效应，即“城市人”和“本地人”。为了进一步识别这两种效应，我们利用流动人口内部户籍差异来进行分析。由于非农户籍的流动人口本身已经具备城镇户口，融入本地只可能通过“本地人”身份而不是通过“城市人”身份起作用，因此，如果农业户籍流动人口与非农户籍流动人口“本地人”这一变量对过度劳动影响存在差异，则说明对农业户籍流动人口“城市人”身份也在发挥作用。我们在回归方程中加入了户口性质 $rural_hukou_i$ 这一二元变量构造模型来识别两种身份在起作用。 $rural_hukou_i$ 取值为 1 表示户口性质为农业户口。利用交互项分离两种效应的回归方程如 (2) 式：

$$overwork_i = \alpha_i + \alpha_1 SocialIdentity_i + \alpha_2 rural_hukou_i + \alpha_3 SocialIdentity_i \times rural_hukou_i + \alpha_0 X_i + \varepsilon_i \quad (2)$$

$SocialIdentity$ 系数反映了具有城镇户籍流动人口在流入地的“本地人”效应，而交互项系数反映了不同户籍流动人口身份认同对过度劳动影响之差即“城市人”身份影响。

（三）变量定义和描述性统计

根据本文使用的调查数据，流动人口的每周平均劳动时间为 56.8 小时，其中 75.15% 的劳动者每周劳动时间超过 40 小时，62.48% 的劳动者每周平均劳动时间超过 50 小时。这一数据表明，流动人口存在着较为严重的过度劳动问题。

下面给出了流动人口劳动时间的频率分布直方图，从图中可以明显看出流动人口劳动时间的分布存在明显的右偏，流动人口劳动时间集中在 30-70 小时。超 40% 流动人口的劳动时间为 50-60 小时，劳动时间超过 60 小时的流动人口所占比例超过五分之一。

^① 本文主要采用了线性概率模型进行估计，使用 probit 模型估计结果完全一致。

^② 为验证结论的稳健性，本文在后文使用了不同的过度劳动时间划分标准进行了检验。

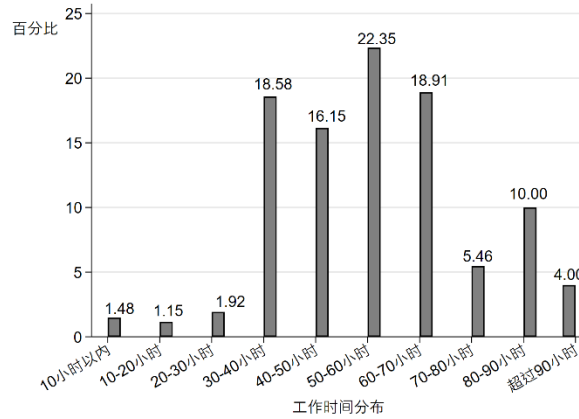


图 流动人口劳动时间分布直方图

从流动人口的职业身份上看，政府机关办事人员以及专业技术人员的每周平均劳动时间最短，均低于 50 小时；而从事经商、餐饮的流动人口每周平均劳动时间超过 60 小时；生产工人、快递人员以及保安的每周平均劳动时间也较高。

另外，我们根据流入地城市将流动人口数据与中国劳动力动态调查（CLDS）2016 年的流入地城镇户籍居民劳动时间数据进行了匹配，比较了城镇户籍人口和流动人口的劳动时间差异，结果在附表 1 中给出。通过差异比较我们发现，流入地城镇户籍人口的平均劳动时间为 43.5 小时，显著低于流动人口。

本文核心解释变量为流动人口的本地人身份认同，流动人口社会身份认同由调查问卷中“认为自己是不是本地人”进行衡量。基于受访者的答案，1 代表流动人口认同本地身份，0 代表不认同。

变量的定义以及描述性统计如表 1 所示。可以看出，具有不同户口性质和社会身份认同的流动人口在过度劳动情况上存在着差异。农业户籍流动人口以及不具有本地身份认同的流动人口发生过度劳动的比例较高。换一个视角看，计算可得认同本地人身份的流动人口有 64.3% 存在过度劳动，而不认同本地人身份的流动人口有 68.0% 存在过度劳动，高出近 4%。市场工资水平可能也是影响过度劳动因素，为了尽量得到外生的市场工资率，本文使用了同一城市同行业相似岗位类别人口收入对时间比的平均作为岗位平均工资率的度量。除此之外，过度劳动人群和无过度劳动人群的其余各项特征都有较大差异，比如男性、低学历群体过度劳动可能性更大。这意味着回归分析控制这些变量十分必要。

表 1 变量定义及描述性统计

变量	无过度劳动		过度劳动		t-test 概率
	样本量	均值	样本量	均值	
本地人身份认同	40157	0.78	75227	0.75	0.000***
户籍身份：农业户籍	40157	0.69	75227	0.84	0.000***
同城相似岗位平均工资率	40157	36.45	75227	16.39	0.000***
性别（男性=1）	40157	0.54	75227	0.60	0.000***
年龄	40157	38.04	75227	39.59	0.000***
婚姻状况（已婚=1）	40157	0.96	75227	0.96	0.026**
健康状况	40157	0.99	75227	0.98	0.000***
党员或共青团员	40157	0.11	75227	0.05	0.000***
本地家庭经济支出（对数）	40157	6.78	75227	6.66	0.000***

老家是否有土地	40157	0.38	75227	0.54	0.000***
家庭子女数量	40157	1.30	75227	1.55	0.000***
工作行业：农业	40157	0.03	75227	0.03	0.089*
工作行业：制造业	40157	0.04	75227	0.01	0.000***
工作行业：服务业	40157	0.59	75227	0.71	0.000***
工作行业：社会工作、 教育和其他	40157	0.11	75227	0.14	0.000***
职业：政府机关办事人员 及专业技术人员	40157	0.20	75227	0.05	0.000***
职业：农民	40157	0.02	75227	0.02	0.035**
职业：商户及服务人员	40157	0.50	75227	0.67	0.000***
职业：工业产品生产者	40157	0.21	75227	0.22	0.012**
职业：其他	40157	0.07	75227	0.04	0.000***
学历：未上过学	40157	0.02	75227	0.03	0.000***
学历：小学	40157	0.10	75227	0.18	0.001***
学历：初中	40157	0.35	75227	0.52	0.000***
学历：高中	40157	0.23	75227	0.20	0.000***
学历：大学专科	40157	0.16	75227	0.06	0.000***
学历：大学本科	40157	0.13	75227	0.02	0.000***
学历：研究生	40157	0.01	75227	0.00	0.000***

注：***、**、*分别代表 1%、5%和 10%水平上显著差异。

四、 统计结果

（一）基础回归结果

表 2 前两列报告了流动人口社会认同情况对过度劳动影响的基准 OLS 估计结果。其中，（1）列报告了全部流动人口的身份认同对是否过度劳动情况的影响；（2）列报告了加入交互项后的回归结果，能够看到对不同户籍流动人口的影响差异。

从 OLS 回归结果中可以看出，流动人口是否认同本地人身份能够显著影响过度劳动。当流动人口具有本地身份认同时，全部流动人口每周工作时间超过 50 小时的概率平均下降 2.8%。农业户籍本身会显著增加过度劳动发生的概率。（2）列的结果表明，在加入交互项进行回归后，本地人身份认同的回归系数仍然显著为负，意味着去除了“城市人”身份后的“本地人”身份仍然对减缓过度劳动有显著作用。去除“城市人”身份的本地身份认同能够缓解过度劳动的原因可能是由于心理认知上发生改变，也可能是出于“向上模仿”机制，即流动人口虽然具有城镇户口，但大多到更大规模更高层级城市工作。认可本地身份的流动人口会以本地人为参照对象调整行为模式，包括劳动行为的改变，心理认同或向上模仿都将最终导致流动人口劳动时间的缩短。但是，农业户籍与本地身份认同的交互项回归系数显著为正。这说明农村户籍流动人口获得“城市人”身份反倒会加剧过度劳动。除户籍身份与本地人身份认同对过度劳动情况的影响外，所处行业工资率对过度劳动的影响同样显著，工资率上升会导致过度劳动可能性下降。这意味着低工资率流动人口倾向于更多的劳动时间，这种收入补偿会提高其绝对和相对收入水平。即使在控制了工资率后，户籍身份与本地身份认同对过度劳动时间仍然存在显著的影响。由于 OLS 模型无法解决可能的内生性问题，我们将采用工具变量方法对此问题进一步检验。

（二）内生性讨论及工具变量法估计

有几个原因导致流动人口的社会身份认同状况可能存在内生性。第一，该指标通过主观变量进行衡量，可能存在测度偏误。由于在调查问卷中，身份认同感相关问题需要受访者在完全不同意、不同意、基本同意和完全同意几个选项中进行选择，受访者可能无法准确地表达出自己的感受，在回答身份认同相关问题时，可能会倾向于选择更为积极的答案。第二，流动人口的劳动时间和社会身份识别变量之间可能存在着反向因果。流动人口过长的劳动时间可能限制了流动人口的社会交往范围和社会活动内容，进而影响流动人口的社会认同。第三是遗漏变量问题：由于户籍制度等原因造成的流动人口制度性隔离，流动人口集中在劳动密集型或低端产业工作，而这一类工作通常具有较低的福利保障水平和较长的劳动时间要求，导致了流动人口的过度劳动。虽然我们控制了行业和工资率，仍然可能存在类似工作要求等遗漏变量。我们也可能遗漏同时影响劳动时间和社会身份认知的流动人口个性特征变量。为解决上述问题，本文引入工具变量法进行了进一步估计。

本文使用是否跨方言区流动作为社会认同工具变量，这借鉴了已有研究的做法。Jaspal (2009) 利用心理学分析方法证明，语言是社会认同的有效记号。使用同一种语言是表现某一社会群体特征的理想工具，并且促进了群体内个体的身份认同 (Jaspal, 2009)。本文根据《中国语言地图集》的分类，将中国汉语方言分为官话、晋语、赣语、徽语、吴语、湘语、客家话、粤语、平语九大方言区，又将官话区细分为东北官话区、北方官话区、北京官话区、胶辽官话区、中原官话区、兰银官话区、江淮官话区和西南官话区。考虑到流动人口调查涵盖了少数民族聚居区，我们在方言区划分中加入蒙古语区和藏缅语系区。若流动人口户籍所在地与流入地所属方言区不同，则“是否跨方言区流动”变量取值为 1，代表流动人口跨方言区进行了流动。

本文还使用了流入本地时间作为本地人身份认同的工具变量。这是因为流入本地时间越长，流动人口越能够适应流入地生活，从而更可能认同自己的本地人身份。

表 2 后两列报告了使用工具变量两阶段回归 (2SLS) 的结果，工具变量的有效性检验同样也在表中汇报，结果通过了外生性检验。

对全部人口的工具变量回归结果与最小二乘回归结果基本一致，即去除“城市人”身份的“本地人”身份认同能够降低发生过度劳动概率，而农业户籍流动人口发生过度劳动概率更高。但是，排除内生性的影响后，表 2 第 (4) 列差分模型中交叉项系数为负但不显著，显示农业户籍流动人口的身份认同对减小过度劳动概率与城镇户籍流动人口差异不大，这和 OLS 结果不同。即“城市人”身份认同并不能够降低过度劳动可能。差分模型中城镇户籍流动人口产生本地身份认同同样能够降低过度劳动可能，意味着主要是“本地人”而非“城市人”身份在发生作用。

控制变量中，越年轻流动人口越可能发生过度劳动，男性更可能过度劳动；已婚身份、子女数量以及家庭经济支出的增加都提高了过度劳动的可能，这可能是由于经济压力增大要求其更多劳动。家庭拥有土地也会增大过度劳动的可能，这可能是由于拥有土地家庭在城市工作是一种短期行为，在城市短时间内更大密度投入工作之后返回农村。从事服务业的流动人口更可能发生过度劳动。学历对过度劳动的影响基本呈现学历越低过度劳动概率更大的趋势。与学历为研究生及以上的流动人口相比，专科及以下学历的流动人口过度劳动状况更为严重，这可能是因为这部分人群在劳动市场上的选择权更少，只能从事一些繁重的工作。学历为大学本科的群体与研究生过度劳动状况差异不大。工资率较低的流动人口过度劳动情况仍然更多，意味着过度劳动是改进低收入流动人口收入地位的手段。

表 1 流动人口社会身份认同对过度劳动影响：OLS 与 IV 估计

变量	(1)	(2)	(3)	(4)
	基础模型	交互项模型	基础模型	交互项模型
	OLS	OLS	IV	IV
本地身份认同	-0.042*** (0.003)	-0.059*** (0.007)	-0.182*** (0.021)	-0.119*** (0.042)
农业户口	0.020*** (0.004)	0.003 (0.007)	0.015*** (0.004)	0.077** (0.036)
本地身份认同* 农业户口		0.021*** (0.008)		-0.077* (0.045)
同城同行 业工资率	-0.006*** (0.000)	-0.006*** (0.000)	-0.007*** (0.000)	-0.007*** (0.000)
是否有土地	0.072*** (0.003)	0.071*** (0.003)	0.072*** (0.003)	0.072*** (0.003)
性别	-0.002*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
年龄	0.011 (0.007)	0.011 (0.007)	0.012* (0.007)	0.012* (0.007)
婚姻状况	0.005 (0.010)	0.006 (0.010)	0.010 (0.010)	0.010 (0.010)
健康状况	-0.009* (0.005)	-0.009* (0.005)	-0.010* (0.005)	-0.010** (0.005)
是否为党员 或共青团员	0.004 (0.002)	0.004 (0.002)	0.004* (0.002)	0.004* (0.002)
家庭经济支出	0.049*** (0.003)	0.049*** (0.003)	0.047*** (0.003)	0.047*** (0.003)
子女数量	0.035*** (0.002)	0.035*** (0.002)	0.033*** (0.002)	0.033*** (0.002)
工作行业：农业	0.030** (0.014)	0.030** (0.014)	0.039*** (0.015)	0.040*** (0.015)

工作行业： 制造业	0.006 (0.009)	0.006 (0.009)	0.018* (0.009)	0.017* (0.009)
工作行业： 服务业	0.048*** (0.003)	0.048*** (0.003)	0.049*** (0.004)	0.049*** (0.004)
学历：未上学	0.346*** (0.020)	0.347*** (0.020)	0.333*** (0.021)	0.331*** (0.021)
学历：小学	0.357*** (0.019)	0.357*** (0.019)	0.348*** (0.019)	0.346*** (0.019)
学历：初中	0.321*** (0.018)	0.322*** (0.018)	0.316*** (0.018)	0.314*** (0.018)
学历：高中	0.230*** (0.018)	0.231*** (0.018)	0.229*** (0.018)	0.227*** (0.018)
学历：专科	0.073*** (0.018)	0.073*** (0.018)	0.077*** (0.018)	0.075*** (0.018)
学历：大学本科	-0.007 (0.018)	-0.007 (0.018)	-0.004 (0.018)	-0.005 (0.018)
职业：政府机关 办事人员及专业 技术人员	-0.032*** (0.007)	-0.032*** (0.007)	-0.036*** (0.007)	-0.036*** (0.007)
职业：农业产 品生产者	0.026 (0.016)	0.026 (0.016)	0.029* (0.016)	0.029* (0.016)
职业：商业及 服务业从业者	0.152*** (0.006)	0.152*** (0.006)	0.151*** (0.006)	0.151*** (0.006)
职业：工业产 品生产者	0.101*** (0.007)	0.101*** (0.007)	0.089*** (0.007)	0.089*** (0.007)
职业：其他	0.320*** (0.029)	0.335*** (0.029)	0.415*** (0.032)	0.362*** (0.044)
常数项	0.297*** (0.029)	0.312*** (0.029)	0.336*** (0.032)	0.285*** (0.044)
样本量	115384	115384	115384	115,84
R 平方	0.147	0.147	0.131	0.08

一阶段 F 值			1560.9 (0)	1589.6 (0)
Sargan-test			2.4942 (-0.1143)	1.1665 (-0.2801)

注：** p<0.01, * p<0.05, * p<0.1。F-test 报告了第一阶段联合检验的 F 值和显著性水平，在 1% 的显著性水平上说明工具变量是强工具变量；过度识别检验报告了 Sargan 统计值和显著性水平，在 10% 的显著性水平上说明工具变量满足外生性。（3）和（4）使用的工具变量为“是否跨方言区流动”以及“流动时间”。

（三）机制检验

流动人口如果具有本地身份认同会降低过度劳动时间，可能有很多机制能够解释这个结果。由于数据所限，我们检验了可能的三个机制：向上流动并模仿、未来不确定性和本地适应。

1. 向上流动模仿机制

根据前文分析，流动人口具有本地身份认同能够显著缓解过度劳动的原因可能在于其心理认知发生改变导致的模仿效应。一般来讲，流动人口都是从相对落后地区进入到相对发达地区，比如从农村进入城镇，从小城镇进入大都市。根据中国劳动力动态调查 2016 年的调查结果，城镇户籍居民的周平均劳动时间为 44.21 小时，而北京、上海、天津等直辖市的户籍居民周平均劳动时间均低于全国平均水平，其中北京市户籍居民的周平均劳动时间为 42.9 小时；东部省份如山东、浙江城镇户籍居民周平均劳动时间也低于全国平均水平。与此相对应的是宁夏、甘肃、广西等人均 GDP 较低、经济发展较慢、居民人均收入较低的省份，城镇户籍居民的劳动时间均高于全国平均水平。根据城市等级划分，一线城市城镇户籍居民的周平均劳动时间为 42.38 小时，而三线城市城镇户籍居民的周平均劳动时间为 44.49 小时。从中国劳动力动态调查（2016）的统计数据可以看出，经济更为发达地区、城市等级更高的城镇户籍居民平均劳动时间低于经济欠发达地区、城市等级较低的城镇户籍居民平均劳动时间。

另一方面，流动人口刚进入较高等级城市时，他们会面临更为激烈的劳动竞争。但从流动人口的职业身份上看，有 27.1% 的流动人口为小商贩，23.3% 的流动人口为餐饮和服务业从业者，另有 17.2% 的流动人口为建筑、运输及其他产业工人，即流动人口大部分从事较为基础的生产和服务类工作；从教育背景上看，接受大学本科及以上教育的流动人口不足总数的 2%。这说明流动人口难以在短期内快速提升劳动质量，而只能改变劳动数量，即他们多半会努力工作、超时劳动，并且越大城市流动的工作时间越长。我们所用的 2017 年卫健委流动人口数据显示：在直辖市流动人口平均劳动时间为 59.8 小时，副省级城市流动人口平均劳动时间为 57.8 小时，而其他城市流动人口平均劳动时间仅为 50.8 小时，越往高级别城市流动过度劳动时间越多，与城镇本地人口劳动时间差别也越大。随着社会融入，身份认同会使得其模仿本地人的经济行为或进一步与收入较高的本地人竞争社会地位。

“向上模仿”或“向上竞争”都可能是影响过度劳动的因素。其中“向上模仿”可能是解释基础回归结果的机制之一，即通过模仿本地人生活方式而降低过度劳动，而且越大城市的流动人口降低效应会越大。本文根据中国城市行政级别划分将城市划分为直辖市、副省级城市和其他城市。根据调查数据，流动人口除了由农村走向城镇之外，也会发生省内跨市、国内跨省的流动。比较流动人口现居城市与其原籍所属城市级别后我们发现，样本中有 43.9% 的流动人口现居城市行政级别高于原籍所属城市，有 53.1% 的流动人口在同级别城市之间流动，而仅有不足 3% 的流动人口由较高级别城市向较低级别城市流动。流出地与流入地城市级别的差异，也意味着流动人口面临挑战的差异，以及产生身份认同后行为差异变动程度的差异。为验证本地身份认同是通过改变流动人口心理认知的“向上模仿”而改变流动人口行为模式，本文将“是否向较高级别城市流动”和“本地身份认同”作为交互项加入回归，检验本地身份认同对于在不同级别城市间发生流动的流动人口过度劳动的影

响。结果如表 3 所示。为排除模型潜在内生性的影响，我们在表中同时报告了 OLS 和工具变量回归结果。

从回归结果中可以看出，本地身份认同无论在 OLS 还是工具变量估计中都仍然显著为负。控制了内生性问题后，向更高级别城市流动产生过度劳动概率更大。这表明在更高级别城市工作更容易导致过度劳动的发生。这可能是由于更高级别的城市具有更激烈的劳动力市场竞争，使得流动人口需要延长劳动时间。

无论是 OLS 还是工具变量估计，身份认同和向高级城市流动的交互项的回归系数都显著为负，说明产生本地身份认同会减弱向更高级别城市流动造成的流动人口过度劳动。产生这一结果的原因在于，更高级别城市竞争更为激烈，流动人口的生活压力与工作压力也可能更大。向更高级别城市发生流动，对于流动人口而言更可能发生过度劳动。在更高级别的城市具有本地身份认同，意味着在心理认知上发生了更大的转变，也更可能导致劳动行为的变化。

未列出的控制变量中，无论是 OLS 还是工具变量估计，农村户籍流动人口都更可能发生过度劳动，其他变量系数与表 2 基本一致。分样本回归显示，无论在非农业户籍还是农业户籍流动人口样本中，交互项的回归结果都显著为负。这一结果证明了“向上模仿”效应。

表 3 本地身份认同对于在不同级别城市间流动的流动人口过度劳动行为影响

变量	(1)	(2)	(3)	(4)	(5)	(6)
	全样本	非农户口	农业户口	全样本	非农户口	农业户口
	OLS	OLS	OLS	IV	IV	IV
本地身份认同	-0.045*** (0.004)	-0.066*** (0.011)	-0.042*** (0.005)	-0.352*** (0.035)	-0.327*** (0.078)	-0.372*** (0.039)
向更高级别 城市流动	-0.008 (0.006)	0.001 (0.015)	-0.007 (0.007)	0.295*** (0.035)	0.260*** (0.078)	0.319*** (0.039)
本地身份认同* 向更高级别 城市流动	-0.058*** (0.005)	-0.087*** (0.014)	-0.052*** (0.006)	-0.301*** (0.028)	-0.309*** (0.067)	-0.310*** (0.031)
其他控制变量	是	是	是	是	是	是
样本量	115384	24276	91108	115384	24276	91108
R-平方	0.151	0.215	0.101	0.113	0.196	0.051

注：*** p<0.01, ** p<0.05, * p<0.1。由于篇幅所限，其余控制变量回归结果未在表格中展示。工具变量估计通过了强工具变量和外生性检验，此处不再报告结果。

2. 未来去向机制

另一个社会身份认同缓解过度劳动的可能原因是过度劳动是一种临时性行为，通常发生在那些不打算长期停留在本地劳动市场的流动人口中，因此社会身份认同通过影响流动人口的长期居留意愿影响了劳动时间。根据 Kahanec & Shields(2010)的研究，不打算在流入地居留的移民更倾向于进行过度劳动。这是因为他们不需要积累流入地人力资本。在对工资率等各种其他可能影响因素进行了控制之后，他们仍然愿意延长工作时长，以便增加积蓄带回家乡。另外，相较于不打算在流入地居留的移民，有居留打算的移民甚至愿意为增加本地人力资本而选择收入较低的工作。为验证这一

机制，我们选取了问卷中流动人口的返乡意愿、居留意愿^①以及落户意愿^②作为被解释变量与本地身份认同进行回归分析，无返乡打算、打算居留以及打算落户时被解释变量赋值为1。OLS和IV两种方法的估计结果如表4所示。

从回归结果中可以看出，具有本地身份认同会显著削弱流动人口的返乡意愿，增强流动人口的居留意愿和落户意愿。这一结果说明了具有本地身份认同可能会影响流动人口的居留打算，进而影响其时间分配的变化。那些打算居留的流动人口可能会将部分时间花费在积累流入地人力资本上，减少劳动时间。对于打算返乡的流动人口，进入流入地劳动市场仅仅是一种临时性行为，他们将时间全部投入在劳动中以期获得更高收入，最终导致劳动时间的延长。

表4 本地身份认同对流动人口返乡和居留意愿的影响

变量	(1)	(2)	(3)	(4)	(5)	(6)
	无返乡打算	打算居留	落户打算	无返乡打算	打算居留	落户打算
	OLS	OLS	OLS	IV	IV	IV
本地身份认同	0.022*** (0.001)	0.084*** (0.003)	0.102*** (0.003)	0.059*** (0.007)	0.134*** (0.021)	0.077*** (0.022)
控制变量	是	是	是	是	是	是
样本量	115384	115384	115384	115384	115384	115384
R-平方	0.013	0.035	0.084	-	0.033	0.084

注：*** p<0.01, ** p<0.05, * p<0.1。由于篇幅所限，其余控制变量回归结果未在表格中展示。此处回归使用的工具变量为“是否跨方言区流动”。工具变量估计通过了强工具变量检验，此处不再报告结果。

为进一步验证流动人口的将来去向能够影响他们的劳动决策，本文进一步将流动人口的将来去向和过度劳动情况进行了回归，回归结果如表5所示。回归结果表明，当流动人口有居留本地的意愿时，他们的劳动时间会明显缩短。这一回归结果进一步证明过度劳动是一种临时性行为，同时也说明本文与Kahanec & Shields(2010)的研究结论一致。

表5 过度劳动与流动人口居留、落户意愿

	(1)	(2)	(3)
变量	过度劳动	过度劳动	过度劳动
不打算返乡	-0.044*** (0.01)		
居留意愿		-0.014*** (0.004)	
落户打算			-0.059*** (0.003)
控制变量	是	是	是
样本量	115384	115384	115384
R-平方	0.149	0.149	0.152

注：*** p<0.01, ** p<0.05, * p<0.1。

^① 问卷中将不打算居留的流动人口分为了打算返乡的流动人口以及打算流动到其他城市的流动人口，因此在这里对打算返乡的流动人口与打算居留的流动人口进行了区分

^② 落户意愿表示流动人口打算将来把户口迁入流入地。

3. 本地适应机制

本地适应机制能更直观解释流动人口身份认同对过度劳动影响,即具有本地身份认同可能会让流动人口拥有更多时间以用于本地人力资本积累,减少过度劳动。为验证这一点,本文将“是否构建起本地社交网络”、“是否按照当地习俗办事”以及“是否适应本地卫生习惯”作为行为适应的代理变量进行分析。对其估计结果如表6所示。

从结果中可以看出,本地身份认同与本地行为适应之间存在着较强的相关性。根据杨菊华(2009)的研究,流动人口身份认同既是流动人口社会融入的指标之一,又是影响其他维度社会融入的重要因素。本文的分析证实产生本地认同能够促进流动人口产生行为适应。

表6 本地身份认同与本地习俗适应性

	(1)	(2)	(3)	(4)	(5)
变量	本地办事习俗	本地卫生习惯	本地社交网络	本地办事习俗	本地卫生习惯
	OLS	OLS	OLS	IV	IV
本地身份认同	0.040*** (0.003)	0.036*** (0.003)	0.167*** (0.003)	0.139*** (0.028)	0.457*** (0.024)
控制变量	是	是	是	是	是
样本量	115384	115384	115384	115384	115384
R-平方	0.028	0.028	0.073	0.02	-

注:*** p<0.01, ** p<0.05, * p<0.1。由于篇幅所限,其余控制变量回归结果未在表格中展示。此处回归使用的工具变量为“是否跨方言区流动”。工具变量估计通过了强工具变量检验,此处不再报告结果。

为解释适应本地生活方式意味着与本地人经济行为接近,进一步说明为何产生本地身份认同的流动人口通常具有相对较短的劳动时间。本文对流动人口过度劳动情况与本地生活方式进行了回归,回归结果如表7所示。回归结果表明,适应本地习俗会明显缓解过度劳动情况,即产生本地身份认同促进了流动人口适应本地生活方式,而与本地人经济行为的接近缩短了流动人口的劳动时间。

表7 过度劳动与本地习俗适应性

	(1)	(2)	(3)
变量	过度劳动	过度劳动	过度劳动
本地办事习俗	-0.018*** -0.003		
本地卫生习惯		-0.011*** -0.003	
本地社交网络			-0.030*** -0.003
控制变量	是	是	是
样本量	115384	115384	115384
R-平方	0.149	0.149	0.15

注:*** p<0.01, ** p<0.05, * p<0.1。

(四) 异质性检验

为进一步检验身份认同可能由于流动人口职业类型以及性别、年龄、教育背景等特征的差异而导致的对过度劳动影响异质性,我们根据流动人口这些特征对流动人口进行了区分,再次进行了回归分析。这一部分的回归结果在附录中给出。

考虑到流动人口的过度劳动问题可能是由于其在流入地劳动力市场上较本地人面临更为明显的劣势，为实现就业，流动人口可能是被迫而非主动选择长时间工作，我们将流动人口的职业类型进行了区分，讨论不同种类工作对应的工作时间灵活性不同是否会造成工作时长差异。从调查样本上看，流动人口的职业身份主要被分为四大类，即有固定雇主的雇员，无固定雇主的雇员（零散工），雇主和自营劳动者。其中雇员占比最大，接近 56%；自营劳动者次之，所占比例为 40.5%；最后是零散工和雇主。从经验上看，零散工和雇主的工作时间自由度最大；自营者虽然也对自己工作时间有自主权，但往往是个体工商户，有固定的营业时间；雇员工作时间则视工作内容而不确定。对不同职业身份工作人群回归结果如附表 2 所示。

回归结果表明对于所有职业身份的子样本而言，控制了内生性后，产生本地身份认同能够有效地改善他们的过度劳动情况。并且，对零散工和雇主的过度劳动缓解效应确实最大。这一结果说明，对于可以自由调整工作时间的流动人口，产生本地身份认同能显著缓解过度劳动，这也说明过度劳动状况受到主观意愿的影响。

除职业身份外，我们还对流动人口的主要职业进行了划分，从调查样本上看，流动人口的主要职业被分为了以下五类：政府机关办事人员及专业技术人员、商户及服务人员、农民、工业产品生产及运输人员和其他职业。其中商户及服务人员占比为 61%，其次是工业产品生产及运输人员，占比 21%。从经验上看，政府机关办事人员及专业技术人员工作时间较为固定，农民对自己的工作时间有更高的自主权，而商业服务业和工业产品生产运输人员的劳动强度往往比较大，工作时间也比较长。对不同主要职业人群回归结果如附表 3 所示。

回归结果表明，对于除农业产品生产人员以外，控制了内生性后，产生本地身份认同能够有效改善他们的过度劳动情况。但对工业产品生产运输人员的影响最大。这一结果表明产生身份认同对工作时间较长、从事行业相对低端的流动人口影响更高。由于原始数据中对主要职业为“其他”的定义比较模糊，在这里我们没有对这一类型进行讨论。但流动人口主要职业引起的差异明显小于职业身份引起的差异。

本文随后检验了男性和女性流动样本产生身份认同对过度劳动情况的影响差异，检验结果在附表 4 中展示。OLS 结果表明，不同性别流动人口的过度劳动情况都会由于产生本地身份认同而得到缓解。但在利用工具变量解决了可能的内生性问题之后，男性的本地身份认同对过度劳动情况的影响不显著。这可能是由于相对于男性，女性工作时间可以更有弹性地调节。

进一步地，本文将流动人口样本按照 20-29 岁、30-39 岁、40-49 岁、50-59 岁以及 60 岁以上年龄组进行了区分，并估计了不同年龄组流动人口产生身份认同对过度劳动情况的影响，结果在附表 5 中展示。结果表明，在排除了内生性可能存在的影响后，产生身份认同能够显著缓解 60 岁以下所有年龄组的过度劳动情况，对 60 岁以上年龄组人群过度劳动情况影响不显著。这可能是由于 60 岁以上流动人口的劳动时间本身较短导致的。

本文将流动人口按照未上学、高中及以下学历、大专学历、大学本科及以上学历进行了区分，并对各子样本再次进行了回归分析，结果在附表 6 中给出。结果表明，产生本地身份认同缓解过度劳动的效应在学历为大学专科及以上的流动人口中最为明显。原因可能在于学历较高的流动人口在选择职业上具有更高的自主性，因此在建立起本地身份认同和本地社交网络之后，更容易找到新工作并调整工作时间。

（五）对过度劳动时间的分析

上文分析了本地城市身份认同对过度劳动广度影响，为进一步研究流动人口产生本地身份认同对缓解其过度劳动时间深度的影响，本文将过度劳动时间作为被解释变量再次进行了分析。在这部分分析中，本文沿用了类似上文分析所用模型：

$$overwork_time_i = \alpha_i + \alpha_1 SocialIdentity_i + \alpha_0 X_i + \varepsilon_i \quad (3)$$

以及

$$overwork_time_i = \alpha_i + \alpha_1 SocialIdentity_i + \alpha_2 rural_hukou_i + \alpha_3 SocialIdentity_i \times rural_hukou_i + \alpha_0 X_i + \varepsilon_i \quad (4)$$

其中 $overwork_time_i$ 代表存在过度劳动的流动人口的劳动时间，即

$$overwork_time_i = \begin{cases} \text{每周劳动时间} - 50, & \text{如果每周劳动时间} > 50 \text{ 小时} \\ 0, & \text{如果每周劳动时间} \leq 50 \text{ 小时} \end{cases}$$

由于此变量属于归并数据，因此我们利用 Tobit 模型估计 (3) (4) 两式，也利用工具变量方法解决潜在的内生性问题，回归结果在附表 7 中给出所示。从结果可以看到，总体上本地身份认同能够降低过度劳动时间。而采用交互模型的回归结果显示交叉项不显著，意味着这更多来自“本地人”效应而非“城市人”效应。分户籍估计结果类似，无论农业还是非农户籍，产生工作所在地身份认同都会显著降低过度劳动时间。

为了进一步讨论收入补偿机制对流动人口过度劳动情况的影响，我们利用工资率对流动人口的劳动时间进行了分位数回归，这一回归结果在附表 8 中展示。回归结果表明，尽管工资率与本地身份认同均能有效缩短劳动时间，但随着劳动时间的延长，同城同行业平均工资率对缩短劳动时间的影响幅度逐渐减小，而产生身份认同对缩短劳动时间的影响幅度逐渐增强。这一回归结果明身份认同独立于收入补偿机制，对流动人口的过度劳动情况造成了影响。

(六) 对过度劳动定义的再次讨论

在前文的讨论中，我们一直将劳动时间大于 50 小时视为过度劳动的定义。为进一步分析结论对过度劳动时间定义的敏感性，我们将劳动时间大于全样本平均劳动时间 56.8 小时以及劳动时间大于 60 小时视为过度劳动定义再次对模型再次进行了回归，回归结果在附表 9 中给出。从回归结果中可以看出，改变过度劳动的定义后，回归结果仍然与前文的回归结果一致。这也从关键变量定义方面证明了上文结论的稳健性。

五、 总结讨论

本文估计了流动人口社会身份认同对其过度劳动情况的影响。研究发现流动人口具有流入地社会身份认同能够有效缓解过度劳动。导致这一现象的原因在于当流动人口具有本地社会身份认同时，他们倾向于以流入地居民作为参照群体，在心理认知上更接近流入地居民，更适应当地习俗，减少未来继续流动可能。这导致了他们行为模式的改变，有效地减小了过度劳动发生概率并缩短过度劳动时间。

由二元户籍制度所带来的制度性隔离，导致了农业户籍流动人口更容易发生过度劳动。对这部分流动人口而言，产生流入地社会身份认同意味着更为明显的心理认知改变，这一改变也更为明显地缓解了他们的过度劳动情况。对于农村户籍流动人口，身份认同的作用可能同时来自“城市人”和“本地人”两种效应，本文证明了“本地人”效应是最主要的。与之相似，由于我国城市间的经济发展与人口规模存在着明显的差异，向不同级别的城市流动，通常也意味着面临不同的挑战。由小城镇流向大城市的劳动者更容易过度劳动，但产生本地身份认同也意味着更为明显的心理认知改变，从而也会更为明显地改善过度劳动。

本文的研究具有一定政策意义。过度劳动严重损害劳动者的生理和心理健康，流动人口严重的过度劳动情况更是反映出了流动人口与流入地城镇居民之间在劳动就业方面的巨大不平等。促进流动人口城市身份认同，不仅有利于推进社会融合，提高流动人口特别是农村转移人口彻底市民化，也有助于提升流动人口的劳动时间合理化，从而保障流动人口的健康水平和福利水平。已有众多关

于流动人口身份认同影响因素的研究都指出,户籍包含的社会福利是促进流动人口融入城市获得认同感的重要因素。这意味着让流动人口在城市能够享受同等福利,有着更多维度正向意义。

参考文献:

- 崔岩, 2012: 《流动人口心理层面的社会融入和身份认同问题研究》, 《社会学研究》第 5 期。
- 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。
- 郭凤鸣 牟林, 2019: 《政府培训能否缓解农民工的过度劳动?——基于直接效应和间接效应的分析》, 《中国劳动关系学院学报》第 4 期。
- 国家统计局, 2016: 《中国劳动统计年鉴 2016》, 中国统计出版社。
- 郭凤鸣 曲俊雪, 2016: 《中国劳动者过度劳动的变动趋势及影响因素分析》, 《劳动经济研究》第 4 期。
- 郭凤鸣 张世伟, 2018: 《最低工资提升对低收入农民工过度劳动的影响》, 《中国人口科学》第 5 期。
- 李君甫 许多, 2018: 《农民工的居住方式与过度劳动研究》, 《重庆社会科学》第 9 期。
- 马小红 段成荣 郭静, 2014: 《四类流动人口的比较研究》, 《中国人口科学》第 5 期。
- 孙文凯, 2017: 中国的户籍制度现状、改革阻力与对策, 《劳动经济研究》第 3 期
- 孙文凯 李晓迪 王乙杰, 2019: 《身份认同对流动人口家庭在流入地消费的影响》, 《南方经济》第 11 期。
- 王静 王欣, 2013: 《进城农民工超时工作的成因与特征研究》, 《统计研究》第 10 期。
- 王湘红 陈坚, 2016: 《社会比较和相对收入对农民工家庭消费的影响——基于 RUMiC 数据的分析》, 《金融研究》第 12 期。
- 王湘红 孙文凯 任继球, 2012: 《相对收入对外出务工的影响: 来自中国农村的证据》, 《世界经济》第 5 期。
- 肖红梅, 2014: 《城市从业者“过劳”的成因分析——基于北京地区的调查数据》, 《人口与经济》第 3 期。
- 张文宏 雷开春, 2009: 《城市新移民社会认同的结构模型》, 《社会学研究》第 4 期。
- 杨菊华, 2009: 《从隔离、选择融入到融合: 流动人口社会融入问题的理论思考》, 《人口研究》第 33 期。
- 杨菊华, 2010: 《流动人口在流入地社会融入的指标体系——基于社会融入理论的进一步研究》, 《人口与经济》第 2 期。
- 杨菊华, 2011: 《城乡分割、经济发展与乡-城流动人口的收入融入研究》, 《人口学刊》第 5 期。
- 张静, 2010: 《社会身份的结构失位问题》, 《社会学研究》第 6 期。
- 中国社会科学院语言研究所 中国社会科学院民族学与人类学研究所 香港城市大学语言资讯科学研究中心, 2012: 《中国语言地图集》, 商务印书馆。
- Afridi, F. et al(2015), "Social identity and inequality: The impact of China's Hukou System", *Journal of Public Economics* 123:17-29.
- Altman, M.(1999), "New estimates of hours of work and real income in Canada from the 1880s to 1930: Long-run trends and workers' preferences", *Review of Income and Wealth* 45(3): 353-372.
- Bluestone, B. & S.J.Rose(1998), "The macroeconomics of work time", *Review of Social Economy* 56(4):425-441.
- Bowles, S. & Y.Park(2005), "Emulation, inequality, and work hours: Was Thorsten Veblen right?", *Economic Journal* 115(507):397-412.
- Caruso, C.C.(2006), "Possible broad impacts of long work hours", *Industrial Health* 44(4):531-536.

- Cha, Y. & K.A.Weeden(2014), "Overwork and the slow convergence in the gender gap in wages", *American Sociological Review* 79(3):457-484.
- Chen, Y. & S.X.Li(2009), "Group identity and social preferences", *American Economic Review* 99(1):431-457.
- Dahlgren, A. et al(2006), "Overtime work and its effects on sleep, sleepiness, cortisol and blood pressure in an experimental field study", *Scandinavian Journal of Work, Environment & Health* 32(4):318-327.
- Drago, R. et al(2005), "The existence and persistence of long work hours", IZA discussion Paper, No. 1720.
- Ellingsen, T. & M.Johannesson(2007), "Paying respect", *Journal of Economic Perspectives* 21(4):135-150.
- Fang, H. & G.C.Loury(2005), " 'Dysfunctional identities' can be rational", *American Economic Review* 95(2):104-111.
- Golden, L.(2009), "A brief history of long work time and the contemporary sources of overwork", *Journal of Business Ethics* 84(2):217-227.
- Jacobs, J.A. & K.Green(1998), "Who are the overworked Americans", *Review of Social Economy* 56(4):442-459.
- Jaspal, R.(2009), "Language and social identity: a psychosocial approach", *Psychtalk* 64:17-20.
- Kahanec, M. & M.P. Shields(2013), "The working hours of immigrants in Germany: Temporary versus permanent", *IZA Journal of Migration and Development* 2(1):1-15.
- Kuhn, P. & F.A.Lozano(2008), "The expanding workweek? Understanding trends in long work hours among U.S. men, 1979-2004", *Journal of Labor Economics* 26(2):311-343.
- Michelacci, C. & J.Pijoan-Mas(2007), "The effects of labor market conditions on working time: The US-EU experience", CEMFI Working Paper, No. 0705.
- Park, A.F. & D.Wang(2010), "Migration and urban poverty and inequality in China", *China Economic Journal* 3(1):49-67.
- Tajfel, H. & J.C.Turner (1979), "An integrative theory of intergroup conflict", in:W.G.Austin & S.Worchel (eds), *The Social Psychology of Intergroup Relations*, Monterey, CA: Brooks/Cole.

附录:

附表 1 分省市户籍居民与流动人口劳动时间差异

	流入地城镇户籍居民			流动人口			T 检验
	样本量	均值	标准差	样本量	均值	标准差	概率
周平均劳动时间 (分省)	1948	43.5	0.35	115384	58.95	0.3	0.000***
周平均劳动时间 (分城市)	1948	41.86	1.04	115384	58.76	0.65	0.000***

注: ***, **, *分别代表 1%、5%和 10%水平上显著差异。

附表 2 不同职业身份流动人口社会身份认同对过度劳动影响: OLS 与 IV 估计

变量	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	雇员	零散工	雇主	自营劳动者	雇员	零散工	雇主	自营劳动者
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
本地身份认同	-0.057*** (0.004)	-0.015 (0.011)	-0.056*** (0.012)	-0.036*** (0.004)	-0.321*** (0.028)	-0.516*** (0.097)	-0.415** (0.201)	-0.108** (0.049)
农业户口	0.019*** (0.005)	-0.005 (0.016)	0.042*** (0.013)	-0.004 (0.005)	0.009* (0.005)	-0.002 (0.018)	0.028* (0.016)	-0.007 (0.006)
控制变量	是	是	是	是	是	是	是	是
样本量	54737	9613	7515	43519	64350	9613	7515	43519
R 平方	0.17	0.053	0.112	0.047	0.103	-	0.009	0.041

注: *** p<0.01, ** p<0.05, * p<0.1。此处使用的工具变量为“是否跨方言区流动”。工具变量估计通过了强工具变量检验, 此处不再报告结果。

附表 3 不同主要职业流动人口社会身份认同对过度劳动影响：OLS 与 IV 估计

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
变量	政府机关 办事人员 及专业技 术人员	农业产品生 产	商业服务业	工业产品生 产	其他	政府机关办 事人员及专 业技术人员	农业产品生 产	商业服务业	工业产品生 产	其他
	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV	IV
本地身 份认同	-0.038*** (0.010)	-0.039*** (0.004)	-0.055** (0.028)	-0.045*** (0.006)	-0.043*** (0.015)	-0.002 (0.061)	-0.223*** (0.040)	-0.225 (0.239)	-0.374*** (0.042)	-0.437*** (0.127)
农业户口	0.018* (0.010)	0.021*** (0.005)	-0.001 (0.037)	0.024*** (0.009)	0.031* (0.018)	0.020* (0.011)	0.012** (0.005)	0.002 (0.038)	0.019** (0.009)	0.012 (0.020)
控制变量	是	是	是	是	是	是	是	是	是	是
样本量	11764	70387	2354	25037	5842	11764	70387	2354	25037	5842
R 平方	0.165	0.100	0.064	0.092	0.144	0.164	0.072	0.023	-	0.041

注：*** p<0.01, ** p<0.05, * p<0.1。工具变量估计通过了强工具变量检验，此处不再报告结果

附表 4 不同性别流动人口社会身份认同对过度劳动影响：OLS 与 IV 估计

变量	(1)	(2)	(3)	(4)
	女性样本	男性样本	女性样本	男性样本
	OLS	OLS	IV	IV
本地身份认同	-0.044*** (0.005)	-0.037*** (0.004)	-0.098* (0.057)	0.052 (0.051)
农业户口	0.022*** (0.006)	0.025*** (0.005)	0.020*** (0.006)	0.028*** (0.005)
控制变量	是	是	是	是
样本量	49003	66381	49003	66381
R 平方	0.160	0.133	0.158	0.126

注：*** p<0.01, ** p<0.05, * p<0.1。工具变量估计通过了强工具变量检验，此处不再报告结果。

附表 5 不同年龄组流动人口社会身份认同对过度劳动影响：OLS 与 IV 估计

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
变量	20-29 岁	30-39 岁	40-49 岁	50-59 岁	60 岁以上	20-29 岁	30-39 岁	40-49 岁	50-59 岁	60 岁以上
	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV	IV
本地身份认同	-0.041*** (0.007)	-0.041*** (0.005)	-0.038*** (0.006)	-0.032*** (0.009)	-0.077*** (0.026)	-0.248*** (0.048)	-0.194*** (0.038)	-0.304*** (0.051)	-0.301*** (0.086)	-0.201 (0.210)
农业户口	0.035*** (0.010)	0.025*** (0.006)	0.019*** (0.007)	0.030*** (0.011)	-0.007 (0.029)	0.025** (0.010)	0.019*** (0.006)	0.011 (0.007)	0.023* (0.012)	-0.010 (0.030)
控制变量	是	是	是	是	是	是	是	是	是	是
样本量	17507	46689	35187	13651	2350	17507	46689	35187	13651	2350
R 平方	0.162	0.181	0.103	0.074	0.057	0.125	0.163	0.046	0.020	0.048

注：*** p<0.01, ** p<0.05, * p<0.1。考虑到年龄与流动时间相关，此处使用的工具变量为“跨方言区流动”。工具变量估计通过了强工具变量检验，此处不再报告结果。

附表 6 不同教育背景流动人口社会身份认同对过度劳动影响：OLS 与 IV 估计

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
变量	未上学	高中及以下学历	大学专科	大学本科及以上	未上学	高中及以下学历	大学专科	大学本科及以上
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
本地身份认同	-0.035* (0.019)	-0.046*** (0.003)	-0.028** (0.012)	-0.050*** (0.012)	-0.206 (0.181)	0.032 (0.040)	-0.497* (0.289)	-0.633** -0.321
农业户口	-0.016 (0.038)	0.040*** (0.004)	0.042*** (0.011)	0.045*** (0.012)	-0.028 (0.041)	0.044*** (0.005)	0.020 (0.018)	0.030* -0.016
控制变量	是	是	是	是	是	是	是	是
样本量	2613	94705	10684	7382	2613	94705	10684	7382
R 平方	0.030	0.055	0.071	0.063	0.001	0.049	-	-

注：*** p<0.01, ** p<0.05, * p<0.1。考虑到教育水平与语言能力相关，此处使用的工具变量为“流动时间”。工具变量估计通过了强工具变量检验，此处不再报告结果。

附表 7 流动人口本地身份认同对过度劳动时间的影响

	(1)	(2)	(3)	(4)	(5)	(6)
变量	全样本	全样本	交互项	交互项	非农业户口	农业户口
	Tobit	IVTobit	Tobit	IVTobit	IVTobit	IVTobit
本地身份认同	-1.493*** (0.088)	-6.699*** (0.910)	-1.638*** (0.211)	-5.539*** (2.034)	-7.100*** (2.111)	-13.159*** (0.964)
农业户口	0.081 (0.106)	0.463*** (0.175)	-0.070 (0.215)	1.391 (1.756)		
本地身份认同*农业户口			0.299 (0.232)	-0.934 (2.166)		
控制变量	是	是	是	是	是	是
样本量	115384	115384	115384	115384	24276	91108

注：*** $p < 0.01$ ，** $p < 0.05$ ，* $p < 0.1$ 。由于篇幅所限，其余控制变量回归结果未在表格中展示。工具变量估计通过了强工具变量和外生性检验，此处不再报告结果。

附表 8 工资率和本地身份认同对劳动时间影响的分位数回归

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	劳动时间	劳动时间	劳动时间	劳动时间	劳动时间	劳动时间	劳动时间	劳动时间	劳动时间
分位数	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
同城同行业工资率	-0.019*** (0.002)	-0.163*** (0.004)	-0.179*** (0.004)	-0.167*** (0.004)	-0.186*** (0.006)	-0.188*** (0.004)	-0.173*** (0.005)	-0.162*** (0.006)	-0.155*** (0.007)
本地身份认同	-0.080 (0.052)	-1.395*** (0.142)	-1.397*** (0.133)	-1.394*** (0.120)	-1.888*** (0.178)	-1.669*** (0.142)	-2.189*** (0.167)	-2.298*** (0.191)	-2.477*** (0.224)
常数项	35.117*** (0.479)	38.692*** (1.301)	40.296*** (1.224)	40.815*** (1.099)	40.228*** (1.634)	40.189*** (1.304)	42.583*** (1.530)	42.046*** (1.753)	52.146*** (2.059)
样本量	115384	115384	115384	115384	115384	115384	115384	115384	115384

注：*** p<0.01, ** p<0.05, * p<0.1。由于篇幅所限，其余控制变量回归结果未在表格中展示。

附表9 流动人口本地身份认同对过度劳动的影响：更改定义后检验结果

		(1)	(2)	(3)	(4)
	变量	基础模型	交互项模型	基础模型	交互项模型
		OLS	OLS	IV	IV
(>56.8 小时)	本地身份认同	-0.046*** (0.003)	-0.051*** (0.008)	-0.151*** (0.021)	-0.119*** (0.045)
	农业户口	0.017*** (0.004)	0.012 (0.008)	0.013*** (0.004)	0.044 (0.039)
	本地身份认同*农业户口		0.006 (0.009)		-0.039 (0.048)
	同城同行业工资率	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)
	常数项	-0.059* (0.031)	-0.055* (0.031)	0.009 (0.034)	-0.018 (0.047)
	样本量	115,384	115,384	115,384	115,384
	R 平方	0.115	0.115	0.107	0.107
	变量	OLS	OLS	IV	IV
(>60 小时)	本地身份认同	-0.046*** (0.003)	-0.051*** (0.008)	-0.150*** (0.021)	-0.121*** (0.045)
	农业户口	0.017*** (0.004)	0.012 (0.008)	0.012*** (0.004)	0.041 (0.039)
	本地身份认同*农业户口		0.006 (0.009)		-0.035 (0.048)
	同城同行业工资率	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)
	常数项	-0.063** (0.031)	-0.059* (0.031)	0.005 (0.034)	-0.020 (0.047)
	样本量	115384	115384	115384	115384
	R 平方	0.115	0.115	0.107	0.107

注：*** p<0.01, ** p<0.05, * p<0.1。由于篇幅所限，其余控制变量回归结果未在表格中展示。