

# 过度劳动对农民工社会参与的 “挤出效应”研究\*

——来自中国流动人口动态监测调查的经验证据

祝仲坤

**摘要：**加快农民工城市社会融入是推进新型城镇化战略的重要任务，社会参与是衡量农民工城市社会融入进程的核心要素之一。本文以过度劳动为切入点，审视农民工社会参与低水平发展状态，并结合国家卫生健康委员会发布的2017年中国流动人口动态监测调查数据，系统考察过度劳动对农民工社会参与的影响。研究表明：①中国农民工的过度劳动现象非常严重，68.10%的农民工周平均劳动时间超过50小时，45.88%的农民工周平均劳动时间超过60小时。②过度劳动会显著降低农民工的社会参与水平，即过度劳动对农民工社会参与存在“挤出效应”，在运用泊松内生处理效应模型控制潜在的内生性偏误，并运用2014年社会融合专项调查数据重新检验后，结论依然稳健。③过度劳动之所以会降低农民工的社会参与水平，可能是因为过度劳动会造成农民工心理压力过大、主观社会地位下降。

**关键词：**过度劳动 社会参与 农民工 内生性偏误 泊松内生处理效应模型

**中图分类号：**F323.6 **文献标识码：**A

## 一、引言

自改革开放以来，数以亿计的农民从农村走向城市，成为城镇地区劳动力市场的中坚力量，为缔造中国经济增长奇迹做出了巨大贡献。国家统计局发布的《2018年农民工监测调查报告》显示，2018年中国农民工总量达到2.88亿人，其中进城农民工1.72亿人，占农民工总量的比例接近60%<sup>①</sup>。然而，

\*本文系国家自然科学基金青年项目“数字鸿沟对农户福利不平等的影响机理与效应研究——基于Sen的可行能力理论”

（项目编号：71903062）、全国统计科学研究重点项目“农村数字鸿沟的统计测度及对农民福利的影响效应研究”（项目编号：2019LZ31）的阶段性研究成果。作者感谢国家卫生健康委员会流动人口服务中心提供的数据支持，感谢匿名审稿专家及编辑老师的宝贵意见。当然，作者文责自负。

<sup>①</sup>参见 [http://www.stats.gov.cn/tjsj/zxfb/201904/t20190429\\_1662268.html](http://www.stats.gov.cn/tjsj/zxfb/201904/t20190429_1662268.html)。

大多数进城农民工仅仅实现了地域转移与非农化，尚未实现从农民到市民的社会身份转变。已有研究和实践进程表明，加快市民化进程，努力使农民工“沉淀”下来、融入城市社会，以主人翁的姿态参与城市社会活动、以新市民的角色与属性推进城市建设，将是一项重要的历史任务。

社会融入，或称之为“社会融合”，是一个动态的、渐进式的、多维度的综合性概念（悦中山等，2012）。袁方、史清华（2013）基于可行能力理论从防护性保障、社会条件、经济条件、精神感受及政治参与等多个维度进行分析，得出农民工的社会融入程度仍处于低水平阶段。杨菊华等多位人口学领域的专家学者则借鉴西方社会融合理论，从政治、经济、文化、社会、心理等多个维度进行分析，同样得出流动人口的社会融入程度偏低的研究结论（参见杨菊华，2015）。社会参与是社会融入的重要维度，是衡量社会融入的关键要素，数据显示流动人口的社会参与也处于低水平阶段。国家卫生健康委员会发布的2017年中国流动人口动态监测调查（China Migrants Dynamic Survey，下文简称“CMDS2017”）数据显示，超过六成的农民工在城市中没有参加过任何公共活动，超过半数的农民工在城市中没有参加过任何社会组织<sup>①</sup>。已有文献一直在强调与论证农民工社会参与水平偏低的事实，但对于农民工社会参与水平过低的原因的解释仍然比较乏力。为此，本文试图从过度劳动这一视角揭开农民工社会参与水平过低的“黑箱”，探究过度劳动对农民工社会参与的影响，并理清其中的作用机制，为提升农民工社会参与水平与社会融入程度提供决策依据。

过度劳动，指的是劳动者一定时期内超时超强度的就业状态（朱玲，2009；郭凤鸣、张世伟，2018）。由于劳动强度难以度量，学术界多从劳动时间角度探讨过度劳动（或称“超时劳动”）问题。近年来，诸如互联网企业“996”工作制、医疗行业“过劳死”等中国劳动者过度劳动问题受到了全社会的广泛关注。人力资源和社会保障部发布的《中国劳动统计年鉴-2018》数据显示，2017年中国城镇就业人员周劳动时间均值为46.2小时，这一数值明显超过《劳动法》周劳动时间44小时的规定。相比于城镇就业人员的过度劳动问题，流动人口尤其是城乡流动人口（农民工）的过度劳动现象要严重得多。CMDS2017数据显示，中国流动人口周劳动时间均值为58.1小时，其中，城镇户籍流动人口周劳动时间均值为53.1小时，而农村户籍流动人口（农民工）周劳动时间均值高达58.8小时，这一数值不仅比城镇流动人口的平均水平高10.7%，更是比城镇就业人员平均水平高27.3%。上面的数据容易引发思考：严重的过度劳动问题是不是农民工社会参与程度偏低、社会融入进程迟滞的诱因呢？为了回答这一问题，本文结合CMDS2017数据，从社会组织参与和社会活动参与两个层面探究过度劳动对农民工社会参与行为的影响，并尝试以心理压力和主观社会地位等因素为切入点剖析其中的作用机制。

有别于已有文献，本文潜在的边际贡献包括：第一，在研究视角上，本文首次尝试从过度劳动视角理解农民工城市社会参与、社会融入的发展困境，这也是对过度劳动产生的负面影响或社会危害在农民工群体中的一次检视。第二，在研究数据与方法应用上，本文结合最新的CMDS数据，通过寻找合理的工具变量，运用泊松内生处理效应模型、Oster（2019）提出的遗漏变量检验方法，比较精确地识别过度劳动对农民工社会参与行为的影响，并利用更适合非线性估计的KHB方法识别中介效应。

<sup>①</sup>关于社会活动及社会组织的具体内容，后文的变量描述部分会做具体说明。

## 二、文献综述

### （一）社会融入与社会参与的相关研究

社会融入亦被称为“社会融合”，是一个典型的舶来品。西方学术领域对移民社会融合问题的关注可以追溯到 20 世纪早期芝加哥学派帕克（Park）与伯吉斯（Burgess）的研究，此后针对如何衡量移民的社会融合，从戈登的“二维”模型、杨格-塔斯的“三维”模型到恩泽格尔的“四维”模型，学者们的研究不断深化、拓展。与国外研究演进历程类似，国内学者在分析农民工融入城市社会过程中，也尝试从政治、经济、社会、文化、心理等多个维度展开研究。

关于社会参与，已有研究认为它指的是结构性社会融入，是社会成员在制度与组织层面参与公共事务的过程（颜玉凡、叶南客，2019）。对于农民工而言，社会参与是农民工个体与城市社会沟通的纽带，是农民工了解城市、适应城市、认同城市、参与城市的权利意识及行为能力的呈现，其本质是农民工与城市社会磨合互动过程中的社会身份转换与重塑，也是归属感与认同感的“外化”（唐有财、侯秋宇，2017）。因此，社会参与往往被作为衡量社会融入中“社会维度”的核心内容，是衡量农民工融入城市程度的关键指标之一（张文宏、雷开春，2008）。

从社会参与<sup>①</sup>的相关文献看，国内外关于移民社会参与内涵的理解存在差异，西方研究一般认为社会参与是民主制度的核心要素，移民社会参与也被认为是践行西方式民主的重要路径（陈钊等，2014）。而在中国的语境之下，移民社会参与指的是城乡迁移农民工的社会参与问题，多数研究将这一群体的社会参与作为衡量其社会融入的一个重要维度（例如杨菊华，2015；悦中山等，2012）。在专门探讨农民工社会参与行为的研究中，陈钊等（2014）指出，户籍对移民的公共意识有负面影响，进而使得他们在公共参与方面比较消极。虽然收入水平或受教育水平更高的居民的确具有更强的公共意识，但并不明显具有更高的公共参与度。唐有财、侯秋宇（2017）通过构建分析框架，从身份及场域视角探讨了流动人口文体类参与、公益类参与和管理类参与的影响因素。赵玉峰、扈新强（2019）在研究流动人口社会参与时，重点关注了少数民族群体。若将视野切换至整个民众的社会参与问题上，胡安宁（2014）重点强调了社会参与对老年人身心健康水平的改善作用。熊易寒（2012）基于上海郊区的调查，聚焦城乡结合部中产阶层的公众参与问题，指出这一群体的公众参与很大程度上由公共服务的“洼地效应”引起。孙三百（2018）探讨了住房产权与社会参与行为的关系，指出完全住房产权显著提高城市居民制度化公众参与的概率，共享产权对制度化公众参与无显著影响，这两类住房产权对城市居民非制度化公众参与均无显著影响。

### （二）过度劳动的相关研究

过度劳动指的是劳动者在一定时期内超出平均劳动时间和劳动强度的就业状态（朱玲，2009；郭凤鸣、张世伟，2018）。早在马克思所著的《1844 年经济学哲学手稿》中就出现了“过度劳动”一词，此后在《资本论》中，马克思对过度劳动进行了更详细的论述，他指出工人的过度劳动使工厂主获得

<sup>①</sup>社会参与与公共参与、公众参与等概念大同小异，本文中不作具体辨析。

额外的利润，对于工厂主而言，这种利润是难以抵挡的诱惑。在西方的工业化进程中，劳动时间一直是早期工人运动的重要主题，19世纪末美国和欧洲工人终于赢得“八小时”工作制度法律层面的认可（朱玲，2009），学术探索也随之蓬勃发展，西方学者开始从医学、心理学、法学等多个领域探讨劳动者的过度劳动问题，并逐步向经济学、管理学及社会学领域拓展。从过度劳动的概念辨析、程度测量，到过度劳动的现状与成因，再到过度劳动产生的危害及缓解机制，已有研究形成了诸多真知灼见。

目前，在过度劳动领域的文献中，针对农民工（移民）群体的研究大多开展的是农民工过度劳动的成因分析。从社会环境、制度设计、技术变革、企业文化、工作性质等外部因素（郭凤鸣、牟林，2019），到农民工个体的生理、心理等内在因素（卢海阳、梁海兵，2016），已有文献进行了大量的探讨。关于过度劳动产生的经济社会影响，已有文献主要聚焦过度劳动与劳动者主观福利的关系。例如，朱玲（2009）、张抗私等（2018）等研究分析了过度劳动对劳动者健康状况的负面影响；韩彦超、潘泽泉（2016）研究指出，劳动时间越长，劳动者的生活满意度越低。潘泽泉、林婷婷（2015）则指出劳动时间对农民工的社会融入有显著的负向影响，而且与户籍等因素存在交互效应。综上可知，过度劳动领域的相关文献除了关注过度劳动的成因之外，在过度劳动产生的影响方面重点关注的是其对农民工主观福利的负面作用，其中潘泽泉、林婷婷（2015）的研究与本文话题很接近，但他们研究的核心是社会融入，而非农民工的社会参与。

### （三）过度劳动如何影响农民工的社会参与

根据经典的劳动供给与时间分配理论可知，对于每个人来讲时间是固定的，可以划分为劳动和闲暇，劳动时间过长必然会“挤出”闲暇时间（加里·贝克尔，1998）。对于农民工而言，多项研究均指出过度劳动越严重，农民工的闲暇时间越少。例如，朱玲（2009）指出，过度劳动会严重挤占农民工的休闲、娱乐乃至学习时间；潘泽泉、林婷婷（2015）也明确指出，日工作时间超过10小时的农民工，读书、看报、上网的时间均显著减少，劳动时间越长，农民工越可能陷入社会交往的“内卷化”，与本地人的社会交往越受到限制，社会融入进程也可能因此迟滞。由此一来，即便农民工有意愿参与社会活动或社会组织，也会因劳动时间过长、闲暇时间过少而“心有余而力不足”。根据上述分析，本文提出如下研究假说：

假说：过度劳动会显著降低农民工的社会参与水平。

除了闲暇时间“挤出”这一直接影响，过度劳动还可能对农民工的社会参与行为产生间接影响。社会参与是意愿与行为的有机结合，过度劳动不仅会导致农民工“力不足”，还可能导致农民工“心无余”，即由于减弱了农民工的社会参与意愿，进而降低他们的社会参与水平。

一方面，过度劳动很可能会加大农民工的心理压力，滋生苦闷、焦虑、烦躁等消极情绪，降低生活满意度，甚至诱发抑郁等心理疾病。朱玲（2009）指出，长期超时劳动会导致“慢性疲劳”，严重影响农民工身心健康。王笑天等（2017）指出劳动时间的增长会增加劳动者的资源消耗，进而影响劳动者的心理健康及生活满意度。张抗私等（2018）基于2014年度中国劳动力动态调查数据的研究也指出，工作时间过长会使劳动者面临情绪焦虑、沮丧的可能性增大，心理健康状况明显下降。李东平等（2018）基于2013年度中国综合社会调查数据对农民工群体开展研究，也证实了劳动时间过长会使

农民工的身心健康产生更多隐忧。进一步看,已有研究证实,心理状态在农民工形成城市归属感、融入城市的过程中发挥着重要作用。陈延秋、金晓彤(2016)指出积极乐观的心理状态对新生代农民工形成城市归属感有积极影响,从而促进他们的社会参与行为。相反,若过度劳动导致农民工的心理压力过大,很可能降低他们的社会参与水平。根据上述分析,本文提出以下推论:

推论 1: 过度劳动导致农民工心理压力过大,进而降低农民工的社会参与水平。

另一方面,除了心理压力视角的间接影响外,对于农民工而言,过度劳动还可能存在一层特殊的含义。如上文所述,农民工大多处于产业工人的底层,从事着繁重的体力劳动,是过度劳动最为严重的群体,他们的劳动权益受到损害的现象比比皆是,在与本地居民的比较中,他们很可能感受到更多的不公平、被歧视、被排斥等,加重了他们内心的自卑感,导致主观社会地位下降。潘泽泉、何倩(2017)和祝仲坤、冷晨昕(2018)的研究均指出,主观社会地位认知是影响农民工城市社会融入状态的重要因素,徐延辉、史敏(2018)进一步指出,较低的社会地位是“相对剥夺”的结果,这种“边缘化”状态将导致农民工难以融入城市社会、形成城市认同,其社会参与意愿因此减弱,社会参与水平也因此降低。根据上述分析,本文提出以下推论:

推论 2: 过度劳动导致农民工的主观社会地位下降,进而降低农民工的社会参与水平。

### 三、数据、变量与模型

#### (一) 数据来源与说明

本文使用的主要数据来自 2017 年国家卫生健康委员会公布的中国流动人口动态监测调查(CMDS2017)<sup>①</sup>。该调查采用 PPS(probability proportionate to size, PPS)抽样方法,调查对象为在流入地居住一个月以上,非本区(县、市)户口的 15~59 周岁的流动人口,调查样本涵盖了全国 31 个省(自治区、直辖市)和新疆生产建设兵团,样本总量为 169989 个。本文关注的是农民工群体,因此仅考虑因务工经商迁移、处于就业状态的农业户籍流动人口,在对各变量的缺失值、异常值等进行处理后,获得了包含 102484 个观测值的基准样本。

CMDS2017 数据主要聚焦于流动人口的人体特征、迁移特征、就业特征、社会保障及生活居住状况等内容,能够分析过度劳动对农民工社会参与的影响。不过,CMDS2017 数据对流动人口社会地位与心理健康等方面的内容关注较少,因此本文使用国家卫生健康委员会 2014 年组织的社会融合与心理专项调查(下文简称“2014 年社会融合调查”)数据,即 2014 年中国流动人口动态监测调查 C 板块数据,更为细致地探究过度劳动对农民工社会参与的影响机制。相较于 CMDS2017,2014 年社会融合调查专门针对北京(朝阳区)、青岛、厦门、嘉兴、深圳、中山、郑州和成都 8 个城市开展,抽样方法与调查对象与 CMDS2017 完全一致,样本总量为 16000 个,经过清理,本文获得观测值 12420 个。

#### (二) 变量选择与描述

1. 被解释变量——社会参与。社会参与主要包括社会组织参与和社会活动参与两个层面,具体指

<sup>①</sup>具体信息详见国家卫生健康委员会流动人口数据平台 <http://www.chinaldrk.org.cn/home/>。

的是农民工社会组织的参与种类和社会活动的参与种类。①社会组织参与。调查员询问受访者“目前在本地是以下哪些组织的成员？”<sup>①</sup>，受访者在工会、志愿者协会、同学会、家乡商会、老乡会和其他选项中进行多项选择。调查显示，参加同学会和老乡会的农民工比例较大，分别为21.52%和23.89%；参加工会和志愿者协会的农民工占比分别为7.70%和5.77%；参加家乡商会的农民工比例较低，仅为3.30%。②社会活动参与。调查员通过询问受访者“2016年以来您是否有以下行为？”来衡量农民工的社会活动参与情况，行为选项主要包括“给所在单位/社区/村提建议或监督单位/社区/村务管理”“通过各种方式向政府有关部门反映情况/提出政策建议”“在网上就国家事务、社会事件等发表评论，参与讨论”“主动参与捐款、无偿献血、志愿者活动等”和“参与党/团组织活动，参加党支部会议”。调查显示，各项活动的农民工参与比例分别为7.21%、3.96%、5.59%、34.60%和2.94%。

2.核心解释变量——过度劳动。过度劳动一般是指较长时期的超时、超强度的就业状态，由于劳动强度和疲劳状态等指标难以量化，本文以劳动时间衡量农民工是否过度劳动，这也是已有研究惯用的衡量方式（例如郭凤鸣、张世伟，2018；Cha and Weeden，2014）。根据《劳动法》第36条和第38条规定，劳动者平均每周工作时间不超过44小时，且用人单位应当保证劳动者每周至少休息1天；第41条进一步规定，由于生产经营需要，经与工会和劳动者协商后可以延长工作时间，一般每日不得超过1小时；因特殊原因需要延长工作时间的，在保障劳动者身体健康的条件下延长工作时间每日不得超过3小时。基于上述考虑，本文以周劳动时间是否超过50小时或60小时作为衡量过度劳动的指标，其中，周劳动时间是否超过50小时作为后文分析的核心指标。以周劳动时间是否超过50小时或60小时衡量劳动者是否过度劳动也是已有文献中惯用的标准（例如郭凤鸣、张世伟，2018；Cha and Weeden，2014）。此外，为了比较周劳动时间不同的农民工在社会参与上的差异，本文还将周劳动时间划分为3类，分别为：50小时以下、50~60小时，60小时以上。

CMDS2017 问卷调查员向受访者询问“您今年五一节前一周是否做过一小时以上有收入的工作？”<sup>②</sup>，若受访者回答“是”，即表明其处于就业状态，调查员进一步询问“这周工作时间为多少小时？”。据此，笔者设置了“周劳动时间”变量。样本数据的统计结果显示，农民工的周劳动时间为58.77小时，其中，80.77%的农民工周劳动时间超过40小时，78.98%的农民工周劳动时间超过《劳动法》规定的44小时，68.10%的农民工周劳动时间超过50小时，45.88%的农民工周劳动时间超过60小时。本文进一步描述了农民工过度劳动的分类特征（见图1）。具体来讲，从性别方面看，70.6%的男性农民工存在过度劳动，女性农民工过度劳动的比例虽略低，但也高达64.7%；从代际方面看，老一代农民工过度劳动的比例（73.1%）明显高于新生代农民工（64.5%）；从受教育程度方面看，接受过高中及以上教育的农民工过度劳动的比例（55.3%）远低于接受高中以下教育的农民工（73.8%）；从就业状态方面看，自雇（自我雇佣）的农民工过度劳动的比例高达82.5%，而受雇农民工过度劳动的比例为58.5%。以上结果表明，农民工的劳动时间普遍过长，过度劳动现象非常严重，其中男性、

<sup>①</sup>CMDS2017 问卷与2014年社会融合调查问卷中的“本地”均指的是农民工的务工地。

<sup>②</sup>CMDS2017 是在2017年5月份开展的调查。

老一代、受教育程度偏低以及自雇农民工的过度劳动现象尤为严重。

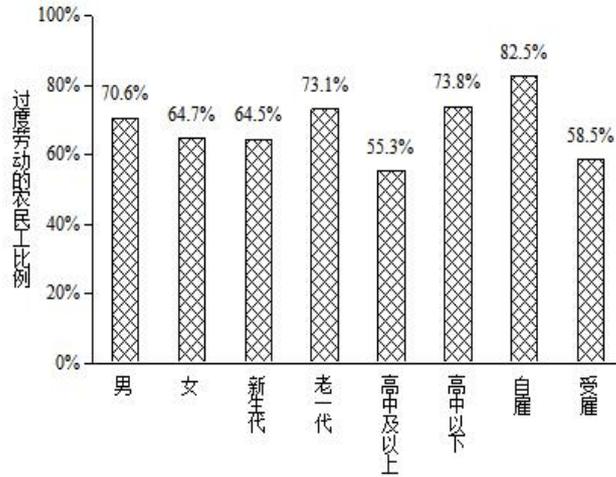


图1 农民工过度劳动的分类特征

为了更为直观地呈现过度劳动与农民工社会参与的关系，笔者绘制了图2。从图中可以看出，周劳动时间在50小时以下的农民工中，至少参与一种社会活动的农民工比例为43.6%，至少参与一种社会组织的农民工比例为49.1%；相比之下，周劳动时间在50~60小时的农民工中，至少参与一种社会活动、至少参与一种社会组织的比例分别为38.3%和43.4%，较周劳动时间50小时以下的农民工分别下降了5.3个百分点和5.7个百分点。进一步看，周劳动时间超过60小时的农民工中，至少参与一种社会活动、至少参与一种社会组织的比例分别为38.2%和39.3%，分别较周劳动时间50小时以下的农民工下降了5.4个百分点和9.8个百分点。以上结果粗略地表明，相比于适度劳动（周劳动时间50小时以下）的农民工，过度劳动（周劳动时间50小时以上）的农民工的社会参与水平更低，换句话说，过度劳动有可能是农民工社会参与水平偏低的诱因之一。

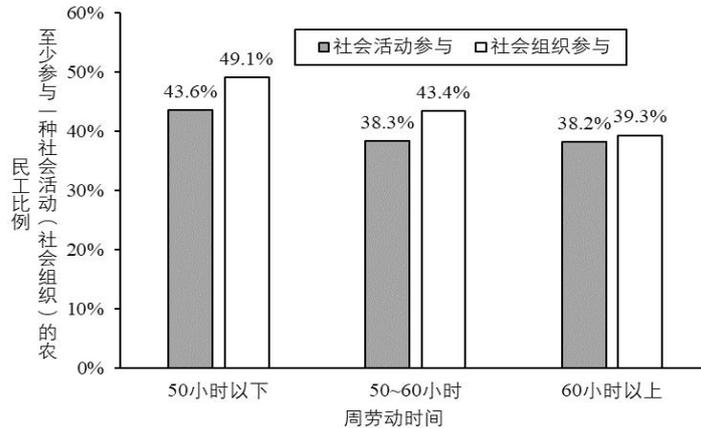


图2 过度劳动与农民工的社会参与

3.控制变量。本文的控制变量包括受访者的年龄、性别、受教育程度、婚姻状态、健康状况、政治身份、家庭成员数量、收入水平、就业身份、住房性质、流动范围、本地居留时间等。需要说明的是，2014年社会融合调查问卷中并未涉及农民工的政治身份与家庭成员数量，另外，在使用CMDS2017

和 2014 年社会融合调查数据进行分析时，健康状况变量的设置是有差异的，表现为：在利用 2014 年社会融合调查数据分析时，健康状况变量被设置为 1~5 的排序变量，即 1 表示健康状况差，5 表示健康状况非常好；而在利用 CMDS2017 数据分析时，健康状况变量被设置为 0-1 变量，1 表示健康，0 表示其他（包括不健康和基本健康）<sup>①</sup>。同样，在使用 CMDS2017 和 2014 年社会融合调查数据进行分析时，收入水平变量的设置也略有差异，表现为：在利用 2014 年社会融合调查数据分析时，收入水平变量为农民工的月工资，而在利用 CMDS2017 数据分析时，收入水平变量为农民工家庭人均月收入。对于收入水平变量，本文在回归分析中均采用自然对数形式。考虑到不同城市的经济发展水平差异明显，本文将农民工所在城市也作为控制变量。最后，本文还控制了农民工所属的行业。值得注意的是，问卷中的行业类别依据《国民经济行业分类》（GB/T4754-2011）共划分为 20 类，但由于篇幅所限，文中并未具体描述，有兴趣的读者可以向作者索要。

变量的含义及其描述性统计见表 1。本文分别汇报了基于 CMDS2017 和 2014 年社会融合调查数据的变量描述性统计结果。

表 1 变量的含义及其描述性统计

变量名称	变量含义和赋值	CMDS2017		2014 年社会融合调查	
		平均值	标准差	平均值	标准差
社会活动参与	在本地参与的社会活动种类（0~5 种）	0.539	0.795	0.647	0.996
社会组织参与	在本地参与的社会组织种类（0~6 种）	0.678	0.937	0.361	0.687
周劳动时间（连续变量）	受访者五一节前一周的劳动小时数（小时）	58.773	17.753	58.574	15.133
周劳动时间（分类变量）					
50 小时以下	50 小时以下=1，其他=0	0.319	0.466	—	—
50~60 小时	50~60 小时=1，其他=0	0.222	0.416	—	—
60 小时以上	60 小时以上=1，其他=0	0.459	0.498	—	—
过度劳动（50 小时标准）	周劳动时间超过 50 小时=1，其他=0	0.681	0.466	0.610	0.488
过度劳动（60 小时标准）	周劳动时间超过 60 小时=1，其他=0	0.459	0.498	0.344	0.475
年龄	受访者 2017 年的年龄（岁）	35.855	9.641	32.383	8.814
年龄平方项	年龄*年龄/100	13.786	7.204	11.264	6.110
性别	受访者性别：男性=1，女性=0	0.569	0.495	0.573	0.495
受教育程度	受访者受教育程度：高中及以上=1，高中以下=0	0.308	0.462	0.347	0.476
受访者婚姻状态					
未婚	未婚=1，其他=0	0.161	0.367	0.280	0.449
在婚	在婚=1，其他=0	0.791	0.407	0.707	0.455

<sup>①</sup>在 CMDS2017 问卷中，健康状况问题实际上包括 4 个选项，即“健康”“基本健康”“不健康，但生活能自理”“生活不能自理”，但本文在对样本异常值的清理中剔除了“生活不能自理”的样本，另外，由于“不健康，但生活能自理”样本仅占总样本的 1.49%，所以本文在分析中将其与“基本健康”的样本进行了合并处理，从而生成了 0-1 变量。

过度劳动对农民工社会参与的“挤出效应”研究

离异或丧偶	离异或丧偶=1, 其他=0	0.049	0.215	0.013	0.114
健康状况	受访者健康状况: 健康=1, 其他=0	0.845	0.362	—	—
	受访者健康状况: 差=1, 一般=2, 好=3, 很好=4, 非常好=5	—	—	3.761	0.971
政治身份	受访者政治身份: 党员或团员=1, 群众=0	0.080	0.271	—	—
家庭成员数量	受访者家庭同住成员的数量	3.177	1.205	—	—
收入水平	受访者家庭人均月收入的自然对数	7.584	0.600	—	—
	受访者月工资的自然对数	—	—	8.068	0.454
就业身份	受访者就业身份: 雇员=1, 自雇佣=0	0.601	0.490	0.687	0.464
受访者住房性质					
租房	租房=1, 其他=0	0.779	0.415	0.928	0.258
保障性住房	保障性住房=1, 其他=0	0.017	0.131	0.007	0.083
自有住房	自有住房=1, 其他=0	0.204	0.403	0.065	0.246
流动范围	受访者流动范围: 跨省流动=1, 省内流动=0	0.511	0.500	0.535	0.499
本地居留时间	受访者到所在城市居住的时间(年)	6.221	5.994	4.011	4.297

注: CMDS2017 的观测值个数为 102484; 2014 年社会融合调查的观测值个数为 12420。

### (三) 计量模型设定

1. 泊松回归与负二项式回归。社会参与属于计数数据 (count data), 这种类型的变量是非负离散型随机变量, 不符合经典线性回归所要求的正态分布, 因此本文考虑采用泊松回归 (Poisson regression) 模型来进行参数估计 (陈强, 2014)。模型设定的具体形式如下:

$$P(y_i | Overwork_i, X_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} (y_i = 1, 2, 3, 4) \quad (1)$$

(1) 式中,  $y_i$  指的是农民工的社会参与,  $Overwork_i$  表示农民工是否过度劳动,  $X_i$  是影响农民工社会参与的各种因素,  $\lambda_i$  指的是“泊松到达率”, 是由各解释变量决定的事件发生次数。

值得注意的是, 泊松分布的局限就在于要求期望和方差相等, 即均等分散 (equal dispersion), 但这一情况可能与实际数据不符。若方差大于期望, 即存在过度分散 (over dispersion), 一般采用负二项式回归 (negative binomial regression) 模型。具体而言, 需要在 (1) 式加入  $\varepsilon_i$  来表征模型的不可观测部分或者个体的异质性 (陈强, 2014)。此时, 模型可以表示为:

$$P(y_i | Overwork_i, X_i, \varepsilon_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} (y_i = 1, 2, 3, 4) \quad (2)$$

关于是否采用泊松回归还是负二项式回归, 可以依据似然比检验 (likelihood ratio test) 进行判别, 似然比检验的原假设为 “ $\alpha = 0$ ”, 即 “不存在过度分散, 应该使用泊松回归”, 若拒绝原假设, 则采用负二项式回归, 若接受原假设则采用泊松回归。

2. 自选择问题。上述回归模型得出一致估计的前提条件为过度劳动是外生变量, 但本文对过度劳

动和适度劳动的农民工样本做均值 t 检验<sup>①</sup>的结果显示, 过度劳动和适度劳动的农民工在多个方面存在显著差异, 这说明过度劳动可能是由个体特征决定的自选择过程。

为此, 本文利用倾向得分匹配法 (PSM) 构造反事实框架进行分析。倾向得分由 Rosenbaum and Rubin (1983) 首次提出, 定义为个体在控制可观测到的混淆变量后, 受到某种解释变量影响的条件概率。具体来讲, 本文旨在比较处理组农民工与相匹配的控制组农民工在社会参与层面的差异, 即得到处理组的平均处理效应 (ATT), 由式 (3) 表示:

$$\tau_{ATT}^{PSM} = E \left\{ E \left[ y_{1i} - y_{0i} \mid \text{Overwork}_i = 1 \right], P(X_i) \right\} \quad (3)$$

(3) 式中,  $y_{1i}$  表示处理组农民工的社会参与水平,  $y_{0i}$  表示与处理组样本相匹配的控制组农民工的社会参与水平,  $P(X_i)$  为倾向得分值,  $X_i$  是可观测到的混淆变量, 与基准回归模型中的控制变量基本一致。虚拟变量  $\text{Overwork}_i = \{0, 1\}$  表示农民工是否过度劳动。

考虑到 PSM 方法在第一阶段存在误设或可观测变量选择不当时, 容易造成估计偏差 (陈强, 2014), 参考 Linden et al. (2016) 的做法, 本文同时运用逆向概率加权法 (IPW) 和逆向概率加权回归调整法 (IPWRA) 进行分析。相比于 PSM 方法, IPW 和 IPWRA 方法可以使处理组和控制组混淆变量的分布更加接近, 所得的 ATT 更具稳健性。两种方法所得的 ATT 分别可以表示为:

$$\tau_{ATT}^{IPW} = \frac{1}{N} \sum_{i=1}^N y_i \frac{\text{Overwork}_i - P(X_i)}{1 - P(X_i)} \quad (4)$$

$$\tau_{ATT}^{IPWRA} = \frac{1}{N} \sum_{i=1}^N y_i \left\{ \left[ \frac{P(X_i)}{1 + P(X_i)} \mid \text{Overwork}_i = 1 \right] \right\} - \left\{ \left[ \frac{P(X_i)}{1 + P(X_i)} \mid \text{Overwork}_i = 0 \right] \right\} \quad (5)$$

3. 内生性问题。除了自选择问题之外, 农民工的脾气秉性、心理状态、过往的经历遭遇等不可观测特征可能同时影响过度劳动与农民工的社会参与, 从而诱发遗漏变量问题。为此, 本文采用工具变量法进行分析。由于核心解释变量为二元离散变量, 被解释变量为计数变量, 采用基于连续变量的两阶段最小二乘法会降低核心解释变量对被解释变量回归结果的有效性 (Angrist and Pischke, 2008), 为此, 本文参考 Terza (1998) 的研究, 运用泊松内生处理效应 (endogenous treatment Poisson regression, ETPR) 模型进行分析。目前, 这种方法在国外得到了比较广泛的应用 (例如 Bratti and Miranda, 2011; Zimmer, 2017)。该模型包括两个阶段方程, 方程的形式如下:

$$\text{Overwork}_i^* = \varphi Z_i + \mu_i, \text{Overwork}_i = \begin{cases} 1, & \text{if } \text{Overwork}_i^* > 0 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

<sup>①</sup>由于篇幅所限, 本文并未展示均值 t 检验结果, 留存备索。

$$f(y_i) = \frac{\exp\{-\exp(\beta Overwork_i + \gamma X_i + \varepsilon_i)\} \{\exp(\beta Overwork_i + \gamma X_i + \varepsilon_i)\}^{y_i}}{y_i!} \quad (7)$$

(6) 式是 ETPR 模型的第一阶段方程, (7) 式是 ETPR 模型的第二阶段方程。其中,  $Overwork_i^*$  是表示农民工是否过度劳动的潜变量, 当  $Overwork_i^* > 0$  时,  $Overwork_i = 1$ , 否则  $Overwork_i = 0$ 。 $Z_i$  是影响农民工过度劳动的控制变量,  $X_i$  是影响农民工社会参与的控制变量。需要注意的是,  $Z_i$  中和  $X_i$  中可以有重叠的变量, 但  $Z_i$  中至少有一个变量不在  $X_i$  中, 以发挥工具变量的作用 (陈强, 2014)。 $\varphi$ 、 $\beta$  和  $\gamma$  是待估参数,  $\mu_i$  和  $\varepsilon_i$  是随机干扰项。

在 ETPR 模型中, 随机干扰项  $(\mu_i, \varepsilon_i)$  服从二元正态分布:

$$\begin{pmatrix} \mu_i \\ \varepsilon_i \end{pmatrix} \sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma\rho \\ \sigma\rho & 1 \end{pmatrix} \right] \quad (8)$$

(8) 式中,  $\mu_i$  的方差被标准化为 1,  $\rho$  是  $(\mu_i, \varepsilon_i)$  的相关系数, 若  $\rho \neq 0$  则说明第二阶段方程存在选择性偏误, 反之, 若  $\rho = 0$  则说明不存在选择性偏误。

ETPR 模型还可以通过计算 ATT 与 PSM、IPW 和 IPWRA 方法得到的 ATT 进行比较, 衡量是否存在不可观测变量带来的选择性偏误, 以及互为因果等其他内生性问题。具体而言, 处理组的  $y_{1i}$  (过度劳动农民工的社会参与水平) 如 (9) 式所示; 控制组的  $y_{0i}$  (适度劳动农民工的社会参与水平) 如 (10) 式所示。

$$E(y_{1i} | Overwork_i = 1) = \exp(\beta_1 Overwork_i + \gamma_1 X_i + \varepsilon_{1i}) \quad (9)$$

$$E(y_{0i} | Overwork_i = 0) = \exp(\beta_0 Overwork_i + \gamma_0 X_i + \varepsilon_{0i}) \quad (10)$$

进一步地, ETPR 模型的 ATT 可以表示为:

$$E(y_{1i} - y_{0i} | Overwork_i = 1) = E \left[ \frac{\{\exp(\beta_1 Overwork_i + \gamma_1 X_i + \varepsilon_{1i}) - \exp(\beta_0 Overwork_i + \gamma_0 X_i + \varepsilon_{0i})\}}{\exp(\frac{\sigma^2}{2}) \frac{\Phi(\sigma\rho + \varphi Z_i)}{\Phi(\varphi Z_i)}} | Overwork_i = 1 \right] \quad (11)$$

为了有效识别 ETPR 模型, 模型中需要加入工具变量。有效的工具变量应当满足以下两点: 第一, 相关性, 工具变量要与农民工是否过度劳动相关。第二, 外生性, 工具变量不能直接影响农民工的社会参与。

本文使用区县层面居住在工作单位或就业场所的农民工比例这一变量作为工具变量。为了避免内生性, 在计算过程中, 笔者剔除了农民工本人。调查员询问了受访者的居住状况, 笔者以样本中农民工本人所在区县中居住在工作单位或就业场所的农民工数量减去农民工本人作为“分子”, 以样本中农民工本人所在区县中农民工总数减去农民工本人作为“分母”, 计算得到本文所用的工具变量。一方面, 居住在工作单位或就业场所的农民工劳动时间可能会更长, 理应符合相关性条件; 另一方面, 农民工

本人所在区县中其他农民工的居住状况不会直接影响农民工本人的社会参与水平，所以也满足外生性条件。同理，本文还使用区县层面过度劳动的农民工比例变量作为工具变量，变量计算方法为：以样本中农民工本人所在区县过度劳动农民工数量（农民工本人除外）作为“分子”，以样本中农民工本人所在区县农民工总数（农民工本人除外）作为“分母”。已有研究表明，来自“分析上层”的工具变量，即在微观个体层面的研究过程中，将县级、社区、村级等层面的解释变量作为工具变量，也是经济学乃至社会科学领域惯用的方法（例如陈云松，2012）。

#### 四、实证结果分析

##### （一）基准回归

本文同时运用泊松回归和负二项式回归模型进行了分析，其中，针对负二项式回归所做的似然比检验在1%的统计水平上显著，拒绝了均等分散的原假设，表明本文应采用负二项式回归，因此下文仅报告负二项式回归的估计结果。

如表2所示，无论是以周劳动时间超过50小时为标准衡量过度劳动，还是以超过60小时为衡量标准，过度劳动都会显著降低农民工的社会参与水平，这初步证实本文的研究假说成立。依据测算的平均边际效应，如果以周劳动时间超过50小时为标准衡量是否存在过度劳动，过度劳动使农民工参与社会活动的预测值下降0.053（回归1），使农民工参与社会组织的预测值下降0.093（回归2）。如果以周劳动时间是否超过60小时为标准衡量是否存在过度劳动，过度劳动使农民工参与社会活动的预测值下降0.031（回归3），使农民工参与社会组织的预测值下降0.090（回归4）。

从控制变量的回归结果看，绝大部分变量也显著影响农民工的社会参与。下面以回归3和回归4的结果为例做简要分析。年龄对农民工的社会参与存在非线性影响，拐点的位置大致在35岁，在35岁之前，随着年龄的增长，农民工社会参与的积极性逐步增强，在35岁附近达到最高点后，随着年龄的增长，农民工社会参与的积极性逐步减弱。相比于女性，男性农民工的社会参与水平显著更高。受教育程度对农民工的社会参与水平具有显著的正向影响，接受过高中及以上教育的农民工的社会参与积极性更强。健康情况越良好，农民工参与社会组织的积极性越强，而参与社会活动的积极性越弱。相比于群众身份的农民工，党员或团员身份的农民工的社会参与积极性更强。家庭成员数量越多，农民工社会参与水平越高。家庭收入水平对农民工社会参与有显著的正向影响。在住房性质方面，相比于租房的农民工，居住在自有住房或保障性住房的农民工社会参与的积极性更强。在流动特征方面，省内流动、在本地居留时间较长的农民工，其社会参与水平较高。总体上看，以上估计结果与以往文献的发现基本一致。不过，考虑到部分控制变量可能存在潜在的内生性问题，还需谨慎解读。

表 2 基准回归——过度劳动对农民工社会参与影响的负二项式回归结果

变量	回归1	回归2	回归3	回归4
	社会活动参与	社会组织参与	社会活动参与	社会组织参与
过度劳动（50小时标准）	-0.053*** (0.006)	-0.093*** (0.006)	—	—

过度劳动对农民工社会参与的“挤出效应”研究

过度劳动 (60小时标准)	—	—	-0.031***	-0.090***
	—	—	(0.005)	(0.006)
年龄	0.044***	0.042***	0.045***	0.043***
	(0.002)	(0.003)	(0.002)	(0.003)
年龄平方	-0.063***	-0.067***	-0.063***	-0.068***
	(0.003)	(0.004)	(0.003)	(0.004)
性别	0.121***	0.134***	0.119***	0.132***
	(0.005)	(0.006)	(0.005)	(0.006)
受教育程度	0.169***	0.186***	0.172***	0.188***
	(0.006)	(0.007)	(0.006)	(0.007)
婚姻状态 (以未婚为对照组)				
在婚	0.018*	-0.016	0.018*	-0.015
	(0.009)	(0.011)	(0.009)	(0.011)
离异或丧偶	0.065***	-0.001	0.065***	-0.001
	(0.015)	(0.017)	(0.015)	(0.017)
健康状况	-0.051***	0.040***	-0.051***	0.039***
	(0.007)	(0.008)	(0.007)	(0.008)
政治身份	0.253***	0.108***	0.254***	0.111***
	(0.008)	(0.010)	(0.008)	(0.010)
家庭成员数量	0.025***	0.024***	0.025***	0.024***
	(0.003)	(0.003)	(0.003)	(0.003)
收入水平	0.101***	0.151***	0.102***	0.152***
	(0.005)	(0.006)	(0.005)	(0.006)
就业身份	0.019***	0.053***	0.021***	0.049***
	(0.006)	(0.007)	(0.006)	(0.008)
住房性质 (以租房为对照组)				
保障性住房	0.125***	0.147***	0.126***	0.148***
	(0.019)	(0.025)	(0.019)	(0.025)
自有住房	0.078***	0.086***	0.079***	0.087***
	(0.007)	(0.008)	(0.007)	(0.008)
流动范围	-0.030***	-0.041***	-0.031***	-0.041***
	(0.006)	(0.007)	(0.006)	(0.007)
本地居留时间	0.007***	0.006***	0.007***	0.006***
	(0.000)	(0.001)	(0.000)	(0.001)
行业类别	已控制	已控制	已控制	已控制
城市虚拟变量	已控制	已控制	已控制	已控制
Wald卡方值	9577.665	10183.084	9475.652	10134.662
Pseudo R <sup>2</sup>	0.041	0.039	0.041	0.039
观测值	102484	102484	102484	102484

注：括号内为稳健标准误，\*\*\*、\*\*、\*分别代表在1%、5%、10%的统计水平上显著。篇幅所限，行业类别、城市虚拟变量的实证结果本文未列出，留存备案。

在表2中基准回归的基础上，笔者进一步基于负二项式回归分析了周劳动时间对农民工社会参与的影响。不仅如此，笔者还将周劳动时间划分为3类，以周劳动时间在50小时以下的农民工为对照组，比较周劳动时间在50~60小时和60小时以上的农民工在社会参与方面表现出的差异。

从表3的结果可以看出，周劳动时间对社会活动参与和社会组织参与的影响在1%的统计水平上显著（回归5和回归6），且系数为负，表明周劳动时间越长，农民工的社会参与水平越低。由回归7和回归8的估计结果可知，相比于周劳动时间在50小时以下的农民工，周劳动时间在50~60小时和60小时以上均会显著降低农民工的社会参与水平，这一结果与表2的估计结果相吻合。进一步地，从平均边际效应看，相比于周劳动时间在50小时以下的农民工，周劳动时间在50~60小时和60小时以上使农民工参与社会活动的预测值分别下降0.052和0.055，使农民工参与社会组织的预测值分别下降0.061和0.117，由此可知，周劳动时间在60小时以上的农民工的社会参与水平更低。

表3 进一步回归——过度劳动对农民工社会参与影响的负二项式回归结果

变量	回归5	回归6	回归7	回归8
	社会活动参与	社会组织参与	社会活动参与	社会组织参与
周劳动时间（连续变量）	-0.001*** (0.000)	-0.003*** (0.000)	— —	— —
周劳动时间（分类变量，以50小时以下为对照组）				
50~60小时	—	—	-0.052*** (0.007)	-0.061*** (0.008)
60小时以上	—	—	-0.055*** (0.006)	-0.117*** (0.007)
控制变量	已控制	已控制	已控制	已控制
Wald卡方值	9500.071	10132.931	9577.771	9514.414
Pseudo R <sup>2</sup>	0.041	0.039	0.041	0.041
观测值	102484	102484	102484	102484

注：括号内为稳健标准误，\*\*\*、\*\*、\*分别代表在1%、5%、10%的统计水平上显著。篇幅所限，控制变量、行业类别、城市虚拟变量的实证结果本文未列出，留存备案。

## （二）纠正选择性偏误

在使用PSM、IPW和IPWRA等方法之前需要进行平衡性检验<sup>①</sup>。本文运用多种手段进行匹配后，各个解释变量的标准化偏误显著下降，表明使用PSM方法可以有效降低处理组及与之匹配的控制组样本的标准化偏误，基本上达到了类似随机实验的效果。

表4报告了基于PSM、IPW和IPWRA方法测算出的ATT结果。从社会活动参与方面看，3种方

<sup>①</sup>篇幅所限，平衡性检验结果没有列出，感兴趣的读者可向作者索要。

法测算的 ATT 分别为-0.051、-0.046 和-0.057，均在 1%的统计水平上显著。从社会组织参与方面看，3 种方法测算的 ATT 分别为-0.081、-0.075 和-0.091，也均在 1%的统计水平上显著。这表明消除了样本间可观测的系统性差异后，过度劳动会显著降低农民工的社会参与水平，本文的研究假说进一步证实。

表 4 PSM、IPW 和 IPWRA 方法的估计结果

变量	估计方法	ATT	标准差	t 统计量
社会活动参与	PSM	-0.051***	0.007	-7.20
	IPW	-0.046***	0.007	-6.65
	IPWRA	-0.057***	0.007	-8.56
社会组织参与	PSM	-0.081***	0.006	-10.56
	IPW	-0.075***	0.008	-9.68
	IPWRA	-0.091***	0.007	-12.28

注：\*\*\*、\*\*、\*分别代表在 1%、5%、10%的统计水平上显著。

### (三) 内生性讨论

1. 基于 ETPR 模型的分析。本文运用 ETPR 模型分析了过度劳动对农民工社会参与的影响，估计结果见表 5。ETPR 模型第一阶段方程的估计结果显示，区县层面居住在工作单位或就业场所的农民工比例和区县层面过度劳动的农民工比例均与农民工过度劳动在 1%的统计水平上显著正相关，表明本文所选的工具变量满足相关性条件。进一步看，社会活动参与和社会组织参与的内生性参数均在 1%的统计水平上显著，且符号为正，表明农民工过度劳动是内生虚拟变量，基准回归结果可能受到不可观测选择性偏误（遗漏变量）的干扰。同时，内生性参数为正表明，基准回归以及 PSM、IPW 与 IPWRA 方法的所得结果，可能因为未能解决不可观测选择性偏误或互为因果等内生性问题而高估了过度劳动对农民工社会参与的回归系数，即低估了过度劳动对农民工社会参与的弱化作用，基于 ETPR 模型估计结果测算出的 ATT 低于（绝对值高于）基于 PSM、IPW 与 IPWRA 方法测算出的 ATT，有力地支撑了这一结论。

表 5 过度劳动对农民工社会参与影响的 ETPR 模型回归结果

变量	回归9	回归10	回归11	回归12	回归13	回归14	回归15	回归16
	过度 劳动	社会活动 参与	过度 劳动	社会组织 参与	过度 劳动	社会活动 参与	过度 劳动	社会组织 参与
	第一阶段	第二阶段	第一阶段	第二阶段	第一阶段	第二阶段	第一阶段	第二阶段
区县层面居住在工作单位 或就业场所的农民工比例	0.144*** (0.052)		0.186*** (0.052)					
区县层面过度劳动的农民 工比例					2.455*** (0.032)		2.452*** (0.031)	
过度劳动		-0.181*** (0.061)		-0.672*** (0.073)		-0.258*** (0.038)		-0.342*** (0.040)
控制变量	已控制	已控制	已控制	已控制	已控制	已控制	已控制	已控制
ATT	-0.099***		-0.193***		-0.147***		-0.252***	

过度劳动对农民工社会参与的“挤出效应”研究

	(0.037)	(0.029)	(0.024)	(0.035)
内生性参数	0.142***	0.611***	0.294***	0.263***
	(0.036)	(0.026)	(0.067)	(0.050)
Wald卡方值	8652.402	8932.698	8675.953	9215.547
观测值	102484	102484	102484	102484

注：括号内为稳健标准误，\*\*\*、\*\*、\*分别代表在1%、5%、10%的统计水平上显著。篇幅所限，控制变量、行业类别、城市虚拟变量的实证结果本文未列出，留存备索。

2. 遗漏变量检验。由遗漏变量导致的估计偏误是考察因果效应的重要阻碍（Angrist and Pischke, 2008），社会科学领域的研究很难穷尽所有的影响因素，遗漏变量似乎难以避免。虽然本文在基准回归分析中尽可能地控制了包括农民工所在城市及行业等在内的多种因素，但仍然无法确保模型中不存在某些不可观测的因素。为了解决遗漏变量带来的偏误，上文的思路是寻找合适的工具变量，并运用ETPR模型进行分析。除此之外，本文尝试从另外一个思路审视遗漏变量问题：假如存在遗漏变量，它会对回归过程产生怎样的影响？估计结果还满足一致性吗？为此，本文依据Oster（2019）提出的两种方法检验潜在的遗漏变量及其对回归过程的影响。

Oster（2019）证明，当回归模型存在不可观测的遗漏变量时，可根据估计量  $\beta^* = \beta^*(R_{max}, \delta)$  获得核心解释变量对被解释变量的一致估计。在本文中， $\beta^*$  衡量的是过度劳动对农民工社会参与的影响， $R_{max}$  指的是若不可观测的遗漏变量能够被观测，回归方程的最大拟合优度； $\delta$  是指选择平衡度，它衡量的是可观测变量、不可观测变量与被解释变量相关关系的强弱（参见马双、赵文博，2019）。具体而言，本文采用两种方法检验过度劳动对农民工社会参与的影响是否因遗漏变量而发生显著变化：方法一，假定  $R_{max}$  是某一数值（一般而言， $R_{max}$  取值为当前回归分析拟合优度的1.3倍），如果  $\beta^* = \beta^*(R_{max}, \delta)$  落在了待估参数  $\beta$  的95%置信区间内，则意味着遗漏变量不会改变过度劳动对农民工社会参与的影响。方法二，与第一种方法类似，同样假定  $R_{max}$  是某一数值，在此基础上，本文计算使  $\beta = 0$  的  $\delta$  的取值，若  $\delta$  取值大于1，也就意味着过度劳动对农民工社会参与的影响不会因遗漏变量而发生显著变化。

由于Oster（2019）的方法主要适用于被解释变量为连续变量情形，因此本部分的分析过程采用OLS回归。如表6所示，过度劳动对农民工社会参与影响的实际计算结果落在了  $\beta$  的95%置信区间内，当不可观测的遗漏变量能够被观测时， $\delta$  的取值大于1，上述结果表明，即便在模型中存在遗漏变量，过度劳动对农民工社会参与的影响也不会因为遗漏变量的存在而发生显著变化，这进一步证实了过度劳动会降低农民工社会参与水平的结论是稳健的。

表6 遗漏变量检验结果

被解释变量	检验方法	判断标准	过度劳动对农民工社会参与影响的实际计算结果	是否通过检验
社会活动参与	方法一	$\beta^* = \beta^*(R_{max}, \delta) \in (-0.120, -0.032)$	-0.056	是
	方法二	$\delta > 1$	2.138	是
社会组织参与	方法一	$\beta^* = \beta^*(R_{max}, \delta) \in (-0.113, -0.088)$	-0.101	是

	方法二	$\delta > 1$	2.592	是
--	-----	--------------	-------	---

注：遗漏变量检验借鉴 Oster (2019) 提出的方法。

3.稳健性检验。由于农民工的社会参与为计数数据，本文采用负二项式回归及 ETPR 模型进行分析。但依据 CMDS2017 数据可知，农民工社会活动参与、社会组织参与的均值仅分别为 0.539 和 0.678，这意味着无论是社会活动，还是社会组织，农民工的参与程度均比较低。为此，本文考虑将被解释变量转化为 0-1 变量<sup>①</sup>，运用普通的 Probit 模型和递归双变量 Probit (recursive bivariate Probit, RBP) 模型进行稳健性检验。目前，这类方法已经得到比较广泛的应用 (例如孙三百, 2018)。

从表 7 的结果可知，无论是运用 Probit 模型还是运用 RBP 模型 (例如选取区县层面居住在工作单位或就业场所的农民工比例作为工具变量) 进行分析，估计结果都显示过度劳动会显著降低农民工的社会参与水平，这一结果与本文的研究假说高度一致，进一步证实了估计结果的稳健性。

表 7 稳健性检验——过度劳动对农民工社会参与影响的 Probit 模型和 RBP 模型回归结果

变量	Probit 模型		RBP 模型	
	回归 17	回归 18	回归 19	回归 20
	社会活动参与	社会组织参与	社会活动参与	社会组织参与
过度劳动	-0.017*** (0.003)	-0.045*** (0.003)	-0.582*** (0.093)	-0.649*** (0.090)
内生性参数			0.331*** (0.061)	0.327*** (0.061)
控制变量	已控制	已控制	已控制	已控制
Wald卡方值	6906.541	7722.673	7577.088	8398.579
观测值	102484	102484	102484	102484

注：括号内为稳健标准误，\*\*\*、\*\*、\*分别代表在 1%、5%、10% 的统计水平上显著。篇幅所限，控制变量、行业类别、城市虚拟变量的实证结果本文未列出，留存备案。

本文还尝试将数据替换为 2014 年社会融合调查数据进行稳健性检验。如表 8 所示，本文基于 2014 年社会融合调查数据，分别运用负二项式回归和 ETPR 模型分析了过度劳动对农民工社会参与的影响。从表 8 不难看出，无论是采用负二项式回归，还是采用 ETPR 模型，估计结果都表明，过度劳动对农民工的社会参与水平有显著的负向影响，研究假说得到进一步证实。

表 8 稳健性检验——基于 2014 年社会融合调查数据的估计结果

变量	负二项式回归		ETPR 模型	
	回归 21	回归 22	回归 23	回归 24
	社会活动参与	社会组织参与	社会活动参与	社会组织参与
过度劳动	-0.092***	-0.047***	-0.099***	-0.048***

<sup>①</sup>若农民工在本地未参与任何社会活动 (或社会组织)，变量取值为 0；若农民工在本地至少参与一种社会活动 (或社会组织)，变量取值为 1。

过度劳动对农民工社会参与的“挤出效应”研究

	(0.016)	(0.013)	(0.018)	(0.013)
内生性参数			0.663***	0.142***
			(0.097)	(0.053)
控制变量	已控制	已控制	已控制	已控制
Wald卡方值	2085.053	1784.191	2020.710	1483.237
观测值	12420	12420	12418	12418

注：括号内为稳健标准误，\*\*\*、\*\*、\*分别代表在1%、5%、10%的统计水平上显著。篇幅所限，控制变量、行业类别、城市虚拟变量的实证结果本文未列出，留存备索。

#### (四) 异质性分析

上文已经得出过度劳动对农民工的社会组织参与存在显著的负向影响，但本文的被解释变量为农民工参与社会组织的种类，因此只能得出过度劳动会减少农民工参与社会组织的种类，无法识别过度劳动对农民工参与不同类型社会组织影响的异质性。本文中的社会组织涉及工会、志愿者协会、同学会、家乡商会、老乡会等多种类型，具有较强的异质性，其中，家乡商会、老乡会等社会组织的家乡属性明显，工会、志愿者协会、同学会具有鲜明的本地标签，更能体现农民工融入本地社会的意愿。接下来，笔者以农民工是否参与工会、志愿者协会、同学会等社会组织为被解释变量，以过度劳动为核心解释变量，通过建立 Probit 模型检验过度劳动对农民工参与不同类型社会组织影响的异质性。考虑到 Probit 模型的估计系数不直观，只能从显著性和参数符号方面给出有限信息，本文汇报过度劳动对农民工社会组织参与的边际效应（陈强，2014）。

如表 9 所示，基于 Probit 模型的估计结果表明，过度劳动对农民工参与工会、同学会和志愿者协会等本地标签比较鲜明的组织的负面作用更为明显，而对参与老乡会、家乡商会等具有明显家乡属性的组织的负面作用相对偏弱。从这层意义上讲，若如上文那样直接将社会组织类型简单加总作为被解释变量做回归分析，那么，过度劳动对农民工融入本地社会的负面影响可能会被低估。这也就意味着，过度劳动加剧农民工社会参与行为的“内卷化”倾向，在过度劳动的重压之下，农民工往往难以摆脱城市的社会排斥、突破自身的“舒适圈”，融入城市的理想固然丰满，现实却也异常骨感。

表 9 过度劳动对农民工参与不同类型社会组织影响的估计结果

变量	回归25	回归26	回归27	回归28	回归29	回归30
	工会	志愿者协会	同学会	家乡商会	老乡会	其他
过度劳动	-0.027***	-0.017***	-0.023***	-0.014***	-0.005***	-0.007***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.002)
控制变量	已控制	已控制	已控制	已控制	已控制	已控制
Wald卡方值	5704.710	2638.152	10198.956	3736.074	1789.818	1690.460
Pseudo R <sup>2</sup>	0.131	0.058	0.108	0.035	0.069	0.034
观测值	102484	102484	102484	102484	102484	102484

注：括号内为稳健标准误，\*\*\*、\*\*、\*分别代表在1%、5%、10%的统计水平上显著。篇幅所限，控制变量、行业类别、城市虚拟变量的实证结果本文未列出，留存备索。

(五) 中介效应分析

为了验证过度劳动对农民工社会参与的间接影响，本文采用中介效应模型进行分析。已有的相关研究大多基于 Baron and Kenny (1986) 提出的分析方法，但考虑到这一方法主要适用于线性模型，为了避免估计结果存在潜在的偏误，本文采用 Karlson、Holm 和 Breen 三位学者于 2010 年提出的适用于非线性模型的 KHB 方法分析过度劳动对农民工社会参与的间接影响 (梁童心等, 2019)。KHB 方法可将核心解释变量对被解释变量的总效应分解为直接效应和间接效应，其中，直接效应指的是核心解释变量不通过任何变量对被解释变量的影响，而间接效应指的是核心解释变量通过中介变量对被解释变量产生的影响。近年来，这一方法得到了比较广泛的应用 (例如石智雷、吴志明, 2018)。

在 2014 年社会融合调查中，调查员向受访者询问了“您觉得您在多大程度上可以选择和掌握自己的生活呢？”，受访者从 1~10 的选项中选择，1 表示根本无法掌握，10 表示完全可以掌握，数值越大，表示程度越高。本文以此作为衡量农民工心理压力的指标。根据样本数据，农民工心理压力的均值为 6.747，过度劳动的农民工心理压力的均值为 6.708，适度劳动的农民工心理压力的均值为 6.784；调查员还向受访者询问“与目前居住地的亲戚、朋友与同事相比，您处在哪个位置？”，受访者从 1~10 的选项中选择，数值越高，表示农民工的主观社会地位越高。根据样本数据，过度劳动的农民工主观社会地位的均值为 5.440，适度劳动的农民工主观社会地位的均值为 5.492。上述结果粗略地表明，相比于适度劳动的农民工，过度劳动的农民工的心理压力更大、主观社会地位更低。

表 10 报告的是基于 KHB 方法<sup>①</sup>的估计结果。结果显示，过度劳动对农民工社会参与的总效应为负，其中过度劳动对农民工社会参与的直接效应占据主导地位，估计结果均在 1% 的统计水平上显著，心理压力和主观社会地位两个中介变量发挥的间接效应居于次要位置，估计结果均在 5% 的统计水平上显著。上述结果表明，过度劳动会降低农民工的社会参与水平，本文的研究假说得到证实；进一步看，过度劳动对农民工社会参与的间接效应不容忽视，过度劳动可能会导致农民工心理压力加大、主观社会地位降低，进而导致农民工的社会参与水平下降，本文的推论 1 和推论 2 也得到证实。

表 10 基于 KHB 方法的过度劳动对农民工社会参与的间接影响

被解释变量	核心解释变量	中介变量	总效应	直接效应	间接效应
社会活动参与	过度劳动	心理压力	-0.142***	-0.139***	-0.003**
		主观社会地位	-0.141***	-0.136***	-0.005**
社会组织参与	过度劳动	心理压力	-0.129***	-0.124***	-0.005**
		主观社会地位	-0.127***	-0.123***	-0.004**

注：\*\*\*、\*\*、\* 分别代表在 1%、5%、10% 的统计水平上显著。篇幅所限，控制变量、行业类别、城市虚拟变量的回归结果未列出，留存备案。

<sup>①</sup>考虑到 KHB 方法更适用于被解释变量为 0-1 变量的情形，本文将社会活动 (或社会组织) 参与设置为 0-1 变量，若农民工未参加社会活动 (或社会组织)，变量取值为 0，若农民工至少参与一种社会活动 (或社会组织)，变量取值为 1。

## 六、结论与启示

### （一）结论

加快农民工城市社会融入进程是重要的历史任务，社会参与是衡量农民工城市社会融入进程的核心要素之一。本文以过度劳动为切入点，审视农民工社会参与低水平的发展状态，并结合 CMDS2017 数据分析过度劳动对农民工社会参与的影响。研究表明：①中国农民工的过度劳动现象非常严重，68.10%的农民工周平均劳动时间超过 50 小时，45.88%的农民工周平均劳动时间超过 60 小时。在农民工群体中，男性、老一代、受教育程度偏低及自我雇佣的农民工的过度劳动现象尤为严重。②过度劳动会显著降低农民工的社会参与水平，这证实了过度劳动对农民工社会参与的“挤出效应”，在运用 PSM、IPW 和 IPWRA 方法纠正选择性偏误，运用 ETPR 模型控制内生性偏误，根据 Oster 的方法衡量遗漏变量产生的影响，并利用 2014 年社会融合调查数据重新检验后，研究结论依然稳健。③异质性分析结果显示，过度劳动对农民工参与本地标签明显的社会组织的负面影响偏大，若直接以社会组织类型简单加总分析过度劳动对农民工社会参与的影响，过度劳动对农民工融入本地社会的负面影响可能会被低估。④过度劳动之所以会降低农民工的社会参与水平，可能是因为过度劳动会使得农民工的心理压力更大、主观社会地位认知更低。

### （二）政策启示

在城乡二元结构下，大量农民进入城市，从事廉价且繁重的劳动，积聚形成“人口红利”，为缔造中国经济增长奇迹做出了重要贡献。然而，长期的过度劳动会严重损害农民工的权益，长此以往，不仅会损害农民工的身心健康，导致农民工“早退”，进而削弱农民工的劳动供给，减弱潜在的人口红利；还可能会导致农民工的社会参与程度偏低、社会融入进程迟滞，甚至激化社会矛盾、形成城市新二元结构。

据此，本文提出，政策部门应当更加重视农民工的过度劳动问题，加快构建缓解农民工过度劳动的政策机制，将缓解农民工过度劳动作为提升农民工社会参与、推动农民工城市社会融入的重要抓手。具体而言，第一，加强规制与监管，尤其是对农民工集中的行业，规范企业单位的劳动时间，尽可能减少非自愿过度劳动，保障农民工的合法权益。第二，强化宣传与引导，提升农民工的维权意识，可以开设微博、微信公众号、法律援助热线等，还可以编制通俗的宣传册，做到线上线下双管齐下，以《劳动法》《劳动合同法》为基础加大宣传普法力度，努力让农民工学会借助法律维护自身合法权益。第三，加强职业辅导与培训，提升农民工的人力资本，尤其针对低教育程度、低技能的农民工，要通过职业辅导与培训手段，努力提升他们的知识与技能水平，提高劳动生产率与竞争力，进而降低他们非自愿过度劳动的概率。第四，要以社区、企业为基本载体，搭建农民工社会参与平台，消除刻板歧视，提升农民工的主观社会地位认知，尤其应注重强化农民工与本地居民的社会互动，拉近两者之间的心理距离，增强农民工的城市归属感与认同感；还应注重农民工的心理疏导，纾解农民工的心理压力，提升其获得感与满足感。

最后需要指出的是，本文还存在一些不足之处。第一，实证分析中，本文以农民工参与社会活动

或社会组织的种类作为被解释变量，可是，对于社会参与行为而言，更为重要的是社会参与的活跃程度，但囿于数据，本文无法获取这一变量。与之类似的是过度劳动变量，本文使用的是周劳动时间，同样受数据限制，本文无法获取农民工的劳动强度信息。此外，由于变量设置层面的限制，CMDS2017数据无法识别过度劳动对农民工社会参与的作用机制，所以本文采用2014年社会融合调查数据进行分析。然而，这套数据的样本仅来自全国8个社会融合试点城市，虽然也采用了PPS抽样，但作用机制的分析结果可能并没有全国范围的普遍意义。总之，变量层面的限制是本文的遗憾之一。第二，由于CMDS数据为截面数据，本文无法捕捉农民工劳动时间及社会参与层面的动态过程，而这很可能导致估计结果出现偏差。今后若有变量设置更为丰富、更具针对性的追踪调查数据，将会更精准地识别过度劳动与农民工社会参与之间的因果效应，从而得到更为严谨和细致的研究结论。第三，CMDS2017问卷中的社会组织具有明显的异质性，不同类型的社会组织发挥的作用差异可能很大，例如，“老乡会”“家乡商会”具有明显的家乡属性，农民工参与这类社会组织的目的可能并非是融入城市社会，因此，本文直接将农民工参与的社会组织种类作为被解释变量可能会低估过度劳动对农民工城市社会融入的负面影响。此外，“工会”在农民工群体中是否发挥着推动农民工融入城市的作用，也有待商榷，需要开展专门性的深入研究。

#### 参考文献

- 1.陈强, 2014:《高级计量经济学及Stata应用(第二版)》,北京:高等教育出版社。
- 2.陈延秋、金晓彤, 2016:《心理资本对新生代农民工社会融入的影响——基于社会距离的中介作用》,《青年研究》第1期。
- 3.陈云松, 2012:《逻辑、想象和诠释:工具变量在社会科学因果推断中的应用》,《社会学研究》第6期。
- 4.陈钊、陆铭、徐轶青, 2014:《移民的呼声——户籍如何影响了公共意识与公共参与》,《社会》第5期。
- 5.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。
- 6.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。
- 7.韩彦超、潘泽泉, 2016:《劳动时间与农民工主观幸福感——基于湖南省农民工三融入调查数据的分析》,《西北人口》第1期。
- 8.胡安宁, 2014:《社会参与、信任类型与精神健康:基于CGSS2005的考察》,《社会科学》第4期。
- 9.加里·贝克尔, 1998:《家庭论》,北京:商务印书馆。
- 10.李东平、卢海阳、文晓丽, 2018:《劳动时间、社会交往与农民工身心健康——基于CGSS2013的实证数据》,《调研世界》第3期。
- 11.卢海阳、梁海兵, 2016:《“城市人”身份认同对农民工劳动供给的影响——基于身份经济学视角》,《南京农业大学学报(社会科学版)》第4期。
- 12.马双、赵文博, 2019:《方言多样性与流动人口收入——基于CHFS的实证研究》,《经济学(季刊)》第1期。
- 13.潘泽泉、何倩, 2017:《居住空间、社会交往和主观地位认知:农民工身份认同研究》,《湖南社会科学》第1期。

- 14.潘泽泉、林婷婷, 2015:《劳动时间、社会交往与农民工的社会融入研究——基于湖南省农民工“三融入”调查的分析》,《中国人口科学》第3期。
- 15.石智雷、吴志明, 2018:《早年不幸对健康不平等的长远影响: 生命历程与双重累积劣势》,《社会学研究》第3期。
- 16.孙三百, 2018:《住房产权、公共服务与公众参与——基于制度化与非制度化视角的比较研究》,《经济研究》第7期。
- 17.唐有财、侯秋宇, 2017:《身份、场域和认同:流动人口的社区参与及其影响机制研究》,《华东理工大学学报(社会科学版)》第3期。
- 18.王笑天、李爱梅、吴伟炯、孙海龙、熊冠星, 2017:《工作时长真的不快乐吗? 异质性视角下工作时长对幸福感的影响》,《心理科学进展》第1期。
- 19.熊易寒, 2012:《从业主福利到公民权利——一个中产阶层移民社区的政治参与》,《社会学研究》第6期。
- 20.徐延辉、史敏, 2018:《社会地位与农民工的定居意愿研究》,《湖南师范大学社会科学学报》第3期。
- 21.颜玉凡、叶南客, 2019:《认同与参与——城市居民的社区公共文化生活的逻辑研究》,《社会学研究》第2期。
- 22.杨菊华, 2015:《中国流动人口的社会融合研究》,《中国社会科学》第2期。
- 23.袁方、史清华, 2013:《不平等之再检验:可行能力和收入不平等与农民工福利》,《管理世界》第10期。
- 24.悦中山、李树苗、费尔德曼, 2012:《农民工社会融合的概念建构与实证分析》,《当代经济科学》第1期。
- 25.张抗私、刘翠花、丁述磊, 2018:《工作时长如何影响城镇职工的健康状况? ——来自中国劳动力动态调查数据的经验分析》,《劳动经济研究》第1期。
- 26.张文宏、雷开春, 2008:《城市新移民社会融合的结构、现状与影响因素分析》,《社会学研究》第5期。
- 27.赵玉峰、扈新强, 2019:《流动人口社会参与的民族差异——基于2014年流动人口动态监测的实证研究》,《西北人口》第2期。
- 28.朱玲, 2009:《农村迁移工人的劳动时间和职业健康》,《中国社会科学》第1期。
- 29.祝仲坤、冷晨昕, 2018:《住房状况、社会地位与农民工的城市身份认同——基于社会融合调查数据的实证分析》,《中国农村观察》第1期。
30. Angrist, J.D., and J. Pischke, 2008, *Mostly Harmless Econometrics: An empiricist's companion*, Princeton: Princeton university press.
31. Baron, R.M., and D.A. Kenny, 1986, "The Moderator-mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations", *Journal of Personality & Social Psychology*, 51(6): 1173-1182.
32. Bratti, M., and A. Miranda, 2011, "Endogenous Treatment Effects for Count Data Models with Endogenous Participation or Sample Selection", *Health Economics*, 20(9): 1090-1109.
33. Cha, Y., and K. Weeden, 2014, "Overwork and the Slow Convergence in the Gender Gap in Wages", *American Sociological Review*, 79(3): 457-484.
34. Linden, A., S. Uysal, A. Ryan, and J. Adams, 2016, "Estimating Causal Effects for Multivalued Treatments: a Comparison of Approaches", *Statistics in Medicine*, 35(4): 534-552.

35.Oster, E., 2019, “Unobservable Selection and Coefficient Stability: Theory and Evidence”, *Journal of Business & Economic Statistics*, 37(2): 187-204.

36.Rosenbaum, P.R., and D.B. Rubin, 1983, “Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome”, *Journal of the Royal Statistical Society*, 45(2): 212-218.

37.Terza, J.V., 1998, “Estimating Count Data Models with Endogenous Switching: Sample Selection and Endogenous Treatment Effects”, *Journal of Econometrics*, 84(1):129-154.

38.Zimmer, D., 2017, “Using Copulas to Estimate the Coefficient of a Binary Endogenous Regressor in a Poisson Regression: Application to the Effect of Insurance on Doctor Visits”, *Health Economics*, 27(3): 545-556.

(作者单位：北京大学国家发展研究院)

(责任编辑：张丽娟)

## **The Crowding-out Effect of Overwork on Migrant Workers' Social Participation: Evidence from China Migrants Dynamic Survey**

Zhu Zhongkun

**Abstract:** This article takes overwork as a starting point, examines the low-level development status of migrant workers' social participation, and carries out a rigorous empirical analysis based on the 2017 China Migrants Dynamic Survey data released by the National Health Commission. The results show that, first of all, the phenomenon of overwork of migrant workers in China is serious, 68.10% of migrant workers work more than 50 hours a week, 45.88% of them work more than 60 hours a week. Secondly, overwork can significantly reduce the level of social participation of migrant workers, that is, overwork has a crowding-out effect on migrant workers' social participation. After using an endogenous treatment Poisson regression (ETPR) model to control potential endogenous bias, the conclusion is still stable. Thirdly, one of the possible reasons why overwork can reduce the social participation level of migrant workers may be that overwork will lead to excessive psychological pressure and the decline of their subjective social status.

**Key Words:** Overwork; Social Participation; Migrant Worker; Endogenous Bias; Endogenous Treatment Poisson Regression Model