

高等教育溢价变动的新趋势与解释

——以流动人口为例

鄢伟波*

内容提要 本文以流动人口为例，研究发现 2011 年以来高等教育溢价呈现持续降低趋势，本科或专科相对于高中学历的年均教育溢价由 2011 年的 8.18% 下降到 2017 年的 4.65%。从供给层面解释高等教育溢价变动，最小二乘估计会因为遗漏变量，低估高等教育供给对溢价的影响。本文利用同一流出地流动人口在流入地集聚的特征，构造工具变量缓解内生性问题，发现特定年龄段流动人口高等教育相对供给每增加 10%，高等教育溢价降低约 1.3% ~ 3%，高等教育供给至少解释了一半以上的高等教育溢价变动。进一步探究高等教育回报的溢出效应，本文发现特大城市即使存在工资溢价，也面临更高的住房租金，高等教育劳动者在不同规模城市的收支比趋于均衡。本文揭示了高等教育溢价变动的新趋势与机理，为劳动者教育投资决策和高等教育政策调整提供了新的启示。

关键词 高等教育供给 教育回报率 流动人口 溢出效应

一 引言

经济学研究者致力于测算真实的教育回报，期望为公众教育投资决策提供成本收

* 鄢伟波，中南财经政法大学财政税务学院，电子邮箱：econotisan@163.com。本文得到教育部人文社会科学研究青年基金项目“新就业形态劳动权益保障研究”（批准号：22YJC790150）、国家自然科学基金青年科学基金项目“新型城镇化进程中户籍改革的劳动力区域再配置效应：动力机制、经济影响与政策优化研究”（批准号：72203028）、中央高校基本科研业务费项目“灵活就业人员社保缴费问题研究：规模测算、生成机制、行为动机与政策优化”（批准号：2722022BY005）的资助。

益分析参考。在政策环境不断变动的背景下，当研究结论与传统观点发生分歧时，政策评估尤其需要在数据、方法和视角上推陈出新，通过良性的争鸣不断接近事实。高校扩招延迟了劳动者进入劳动力市场的时间，不仅有利于缓解过高的失业率压力，还直接提高了劳动者受教育水平。“十三五”时期中国教育公平和质量迅速提升，高等教育进入普及化阶段。高等教育持续扩招下，高等教育溢价发生了怎样的变动？背后的机理是什么？

尽管部分基于横截面数据的研究发现，高校扩招在短期内对大学毕业生就业和收入产生了负向冲击（吴要武、赵泉，2010；邢春冰、李实，2011），但将时间维度拉长来看，基于 2009 年以前数据的测算结果表明，尽管高校扩招后高等教育溢价的增速放缓了，但高等教育溢价整体仍处于上升趋势（Zhang et al., 2005；Meng, 2012；Meng et al., 2013；Li et al., 2017；Zhang, 2017；Feng & Xia, 2022；马光荣等，2017）。受扩招影响的群体在 2003 年以后才开始涌入劳动力市场，这一时期高等教育定价机制的转变释放了部分制度红利，加入世贸组织产生的外部需求强劲。此外，2009 年以前的扩招还吸收了 1981 - 1991 年婴儿潮产生的劳动力供给，农村人口向城市的大幅转移也使得高等教育劳动者比重提升幅度较为缓慢，以上众多因素可能是扩招后高等教育溢价未大幅降低的原因。但 2009 年以后的制度环境变了，农村未受过高等教育的劳动力向城市转移速度放缓，外部需求也由于 2008 年全球金融危机影响而萎靡不振，且 2009 年以后的高校扩招没有婴儿潮作为支撑，而高校扩招的步伐却并未放缓。因此，亟需检验高等教育溢价在 2009 年以后的变动情况并进行解释。

本文使用 2011 - 2017 年中国流动人口动态监测调查数据（CMDS），发现流动人口高等教育溢价逐年降低，本科或专科相对于高中的年均教育溢价由 2011 年的 8.18% 降低到 2017 年的 4.65%。为缓解遗漏变量导致的内生性问题，本文基于 1999 年高校扩招的准自然实验构造工具变量（IV），发现当非高等教育仅包含高中时，最小二乘（OLS）估计结果与工具变量两阶段最小二乘（2SLS）估计结果十分接近。而当非高等教育包含初中时，传统文献中 2SLS 估计结果高于 OLS 估计结果的现象出现，这意味着工具变量识别的局部平均处理效应（LATE）对应的“政策遵从者”（Complier）教育回报率更高^①。探究流动人口高等教育相对供给与高等教育溢价之间的关系，若不考虑不同年龄段劳动者不完全替代性，可能观测到流动人口高等教育相对供给越高的城市，高等教育溢价反而越高的相关性现象。在考虑不同年龄段劳动者不完全替代性后，高

① 即高校（不）扩招，（不）上大学的群体，以下简称“政策遵从者”。

等教育溢价与相对供给负相关，但仍可能因为遗漏变量问题向上误偏。本文参照 Card (2001, 2009) 的研究，基于流动人口流入地聚集特征构造工具变量尝试解决内生性，发现流动人口高等教育相对供给每提升 10%，高等教育溢价降低约 1.3%，当采用省份作为地理单元时降低约 3%，2011—2017 年流动人口教育溢价降低的 53% 以上可由高等教育相对供给增加解释。

本文还尝试探究高等教育可能存在的溢出效应。使用城市层面高等教育相对供给和城市人口规模捕捉溢出效应的来源，在 OLS 模型中引入个人高等教育虚拟变量与规模变量的交互项，发现高等教育劳动者在高等教育相对供给（或城市人口规模）更高的城市教育回报率也更高。将规模变量十等分，本文发现流动人口高等教育溢价随规模提升的现象仅存在于规模前 20% 城市中，而这些城市的住房租金也更高。比较不同学历劳动者的收入支出比，本文发现不同规模城市高等教育与非高等教育劳动者的收支比几乎一致。这意味着即使高等教育溢出效应存在，也伴随着更高的住房租金支出，不同学历流动人口在不同规模城市生活的收支比趋于均衡。

本文的边际贡献有三点：第一，现有关于高等教育溢价变动趋势的研究主要集中在 2009 年以前，得出高校扩招并未显著降低高等教育溢价的结论，这一结论不适合直接移植到当前的政策决策中。尽管部分研究采用 2009 年后的横截面数据测算高等教育溢价（刘生龙等，2016），但尚无研究测算 2009 年后高等教育溢价变动趋势。本文发现高等教育溢价变动的新趋势，延伸了对教育溢价变动的认知和理解。第二，本文利用不同年龄段—学历劳动者不完全替代框架，指出不区分年龄段可能出现高等教育相对供给与高等教育溢价正相关的“伪”回归现象，通过流动人口流入地聚集特征构建的工具变量较好地解决了流动人口学历结构内生性问题。本文测算出不同受教育程度劳动者的替代弹性，为户籍制度调整对收入不平等的影响以及高校扩招政策的成本收益分析提供了重要的参数支撑。第三，本文从住房租金和收支比等维度拓宽了溢出效应的研究视角，发现不同学历流动人口在不同规模城市的收支比趋于均衡，为流动人口选址决策提供了参考。本文以下部分安排如下：第二部分回顾了相关文献，第三部分测算流动人口高等教育溢价并刻画其变动趋势，第四部分解释流动人口高等教育溢价变动并讨论了可能的溢出效应，最后是结论与启示。

二 文献述评

理解收入不平等的传统框架依赖于供给—需求—制度（Supply, Demand, and

Institutions, S-D-I)^①。高等教育溢价的研究文献中,多数将需求视为外生给定或通过工具变量进行控制,来探究高等教育供给变动对教育溢价的影响(Card & Lemieux, 2001; Carneiro & Lee, 2011)。高校扩招可通过数量机制和质量机制对高等教育溢价产生影响。数量机制指高校扩招直接增加了高等教育劳动者供给,在其他条件不变的情况下,高等教育溢价降低。质量机制是指高校扩招导致高等教育质量降低,进而降低高等教育溢价^②。Duflo (2004) 发现,印度尼西亚受过小学教育劳动者相对于未受过小学教育劳动者的比例每增加 10%,小学教育回报降低 4% ~ 10%^③。Bianchi (2020) 则发现,意大利教育扩张中 STEM 专业的大学生增加了 8%,导致扩招前已经上大学的毕业生收入降低 3% ~ 15.5%。陈技伟和冯帅章(2022)认为,高校扩招导致高中毕业生进入大学的概率提升,降低了中等职业教育的生源质量,进而导致中等职业教育相对于高中的教育溢价在高校扩招后消失了,为高校扩招影响教育溢价的质量机制提供了一些启发性证据。Carneiro & Lee (2011) 开创性地对高校扩招降低大学教育溢价的数量机制和质量机制进行了分离,发现美国 1960 - 2000 年高等教育质量的降低导致大学溢价降低了 6%,而这一时期大学溢价仅增加了 20%,表明大学教育溢价变动的质量机制占有重要比例。

高校扩招最直接的影响是高等教育劳动者供给增加了,以上研究通过控制需求来研究供给对高等教育溢价的影响。若考虑需求侧因素,高等教育劳动者供给增加不一定会导致高等教育溢价降低,如技术进步偏向(Acemoglu, 1998; Card & DiNardo, 2002)和高等教育溢出效应(Acemoglu & Angrist, 2000; Moretti, 2004; Li et al., 2017; Feng & Xia, 2022)等因素均可能提升高等教育溢价。对高等教育溢出效应有两类解释:第一类为个体在高等教育劳动者比例高的地区通过信息交流、模仿和干中学等非价格机制提高生产率,可归为非货币外部性(Moretti, 2004; Chen et al., 2020; Bianchi, 2020)。总的来说,这些作用机制并未通过价格机制直接得到补偿,因此可归

① 代表性非传统的分析框架有基于工作任务的职业选择模型(Acemoglu & Autor, 2011),最新研究发现企业对收入不平等具有重要影响(Card et al., 2013; Song et al., 2019),可归为 S-D-I 框架,企业嬗变为一种制度。

② 高等教育质量降低的具体原因可分为两类:一类是不同个体的高等教育回报具有异质性,高校扩招导致教育回报率较低的个体进入大学,使得整体回报率降低;另一类是大规模的高校扩招稀释了人均教育资源,导致高等教育质量降低(Carneiro & Lee, 2011)。分离这两类机制具有重要的政策内涵,陈技伟和冯帅章(2022)讨论了这两类机制, Camacho et al.(2017)进行了分解尝试。

③ 尽管 Duflo (2004) 研究的是小学扩张,但与高校扩张的研究范式无实质性区别。

为非货币外部性。第二类是企业在高等教育劳动者比重更高的地区会加大投资和技术创新，密集的人力资本供给引致的企业投资增加了劳动者收入，可归为货币外部性（Acemoglu & Angrist, 2000, Duflo, 2004; Feng & Xia, 2022）。例如，在 Feng & Xia (2022) 的研究中，高等教育劳动力的供给促进了企业研发和投资活动^①，进而通过技能互补对高学历劳动者产生了更多需求。

尽管高等教育溢出效应的研究构建了丰富的理论模型和逻辑推导，但实证研究还较为欠缺，原因在于严谨识别溢出效应需要直面计量经济学的质疑（Angrist, 2014）。个体教育和外部教育都是内生的，地区不可观测因素可能同时影响地区人力资本以及生产率和工资^②。Acemoglu & Angrist (2000) 通过两个工具变量解决个体和地区层面教育的内生性，对“甚嚣尘上”的教育溢出效应（Romer, 1990）提出了质疑：溢出效应即使存在，经济意义上的显著性也不足以解释不同国家的收入差距^③。Moretti (2004) 是为数不多为教育溢出效应呈现令人信服证据的研究，他展示的两条证据有助于消除对“伪”回归问题的担忧：同一城市内部不同行业间的溢出效应随“经济”距离递减，实物资本不存在与人力资本类似的溢出效应^④。Ciccone & Peri (2006) 指出，能否观测到溢出效应取决于对劳动力可替代性的假设、模型设定和工具变量选取。对知识溢出效应的探索及其在经济增长中的贡献仍然是经久不衰的话题。Akcigit et al. (2018) 利用发明家在研发团队、企业和地区间的交互尝试识别信息传递、知识溢出和干中学等机制，发现高（低）质量的交互有利于提升（降低）发明家的创新质量，且频率更低的交互对应的溢出效应也更小。

在中国溢出效应的研究中，城市规模和密度被认为能提升就业概率（陆铭等，2012）、带来工资溢价（陈飞、苏章杰，2021）和促进服务业发展（钟粤俊等，2020）；胡雯和张锦华（2021）则发现农民工工资与城市密度、迁移距离之间存在非线性关系。对于内生性问题，1953年第一次全国人口普查的城市规模（陆铭等，2012；陈飞、苏章杰，2021）、1933年是否通铁路和1984年人口密度（胡雯、张锦华，2021）、城市土地供应增长率（钟粤俊等，2020）和20世纪50年代初院系调整中每个城市迁入系数

① Che & Zhang (2018) 也发现，高校扩招可能促进企业采用新技术、增加研发投入、雇用高学历劳动者、加快资本深化，进而提升企业全要素生产率。Ma (2020) 也指出，高校扩招提升了企业研发强度和出口升级。

② 相当于同伴效应研究中的内生性效应（Angrist, 2014）。

③ 教育私人回报在10%左右，溢出效应需要达到25%~30%才足以解释不同国家的收入差距。

④ 梁文泉和陆铭（2016）检验了中国行业内企业间、服务业行业间和制造业的人力资本外部性。

量（梁文泉、陆铭，2016），被用来当作城市人口规模和技能结构的工具变量，由于工具变量与断点回归是完全等同的（Vytlacil，2002），对工具变量适用性的严肃论证还有较大改进空间。

针对中国高等教育溢价变动趋势的研究，Meng（2012）和 Meng et al.（2013）发现，2000 - 2009 年间的高等教育溢价增长幅度减缓，但总的来说仍处于增加趋势。马光荣等（2017）意识到高校扩招可能对高等教育溢价产生的影响，利用 Card & Lemieux（2001）提出的不同年龄段、不同学历劳动者不完全替代框架，从供给侧角度讨论了高校扩招导致的高等教育劳动者供给增加对高等教育溢价可能产生的影响^①。对中国高校扩招后教育溢价未显著降低的解释，Li et al.（2017）认为高等教育青年劳动者与中年劳动者技能互补，高校扩招降低青年劳动者教育溢价的同时也提升了中年劳动者教育溢价，最终导致扩招后高等教育溢价并未显著降低。Feng & Xia（2022）通过工业企业数据实证研究发现，高校扩招导致具有技术进步偏向的企业在投资和研发上进行内生调整，避免了高等教育溢价的降低。周茂等（2019）、方森辉和毛其淋（2021）的研究也表明，高校扩招可能提升企业出口质量、推动企业出口升级，加入世贸组织带来的需求侧效应可能恰好为“人力资本扩张”提供了应用场景。

除需求侧因素外，探究高校扩招以来中国高等教育溢价的变动，还须考虑两类典型中国情景：第一，1990 年代中期以来大量未接受过高等教育的农村劳动力开始向城市迁移，在 2004 - 2010 年期间越过“刘易斯拐点”后这一趋势放缓（蔡昉，2018），城市地区用工荒问题开始显现，高校扩招后高等教育溢价变动受到人口转移相对速率的影响。第二，随着计划经济向市场经济转型，劳动力市场定价机制发生了变化，高等教育要素回报的定价机制逐渐转型，2000 年以后大学毕业“包分配”制度转向“自主择业”，劳动力市场定价机制的转变可能释放了被压抑的高等教育溢价（Meng et al.，2013）。

综上，对高校扩招后高等教育溢价变动的研究还较为欠缺，亟需通过 2009 年以后的数据检验高等教育溢价变动。严谨证实高等教育溢出效应的实证研究不足，逻辑与实证互相论证无论对发现新机制还是纠正谬误都是必要的研究范式，而目前后者滞后于前者。与溢出效应有关的研究主要集中于收入，可能忽视了成本。

① 其文显示 1997 - 2009 年间高等教育溢价仍处于上升阶段。

三 流动人口高等教育溢价变动新趋势：2011 - 2017

(一) 计量模型

测算高等教育溢价最基本和重要的步骤是解决内生性问题，Card (1999) 对教育回报中的内生性问题及解决办法做了经典的综述。一般认为不可观测能力等因素与受教育程度存在相关性。对于遗漏变量导致的内生性问题，本文参照刘生龙和胡鞍钢 (2018) 的研究，以 1999 年的高校扩招作为冲击，使用出生季度这一工具变量尝试解决。吴要武 (2010) 论证该工具变量是一个强工具变量，张春泥和谢宇 (2017) 利用中国数据为该工具变量与模糊断点回归之间的关系提供了“第一阶段” (First Stage) 证据^①。具体而言，本文构建如下工具变量模型来测算高等教育溢价：

$$S_i = \lambda + \pi Z_i + X_i' \beta + \eta_i \quad (1)$$

$$\ln wage_i = \alpha + \gamma S_i + X_i' \beta + u_i \quad (2)$$

其中式 (1) 和 (2) 分别为工具变量两阶段最小二乘法的第一阶段和第二阶段估计， S_i 为是否上大学的虚拟变量，取值 1 表示大学本科或专科，取值 0 表示高中或初中， Z_i 为工具变量，对应本文是否受到扩招影响 (*Expan*)， X_i 为一系列控制变量，包括性别、工作经验及其平方项、民族、婚姻状况、户口类型、省份和年份固定效应， $\ln wage_i$ 为对数形式的收入， η_i 和 u_i 为扰动项。

(二) 数据与变量

本文主要使用 2011 - 2017 年中国流动人口动态监测调查数据 (CMDS)，CMDS 是样本量较大、质量较高的连续年份横截面数据。在构造工具变量的具体过程中，按照出生季度对样本进行划分。高等教育的大规模扩招出现在 1999 年 9 月，以 1981 年 6 月 - 8 月出生的群体为参照组，按出生季度设置虚拟变量：1981 年 9 月 - 11 月出生的群体设置为 1，1981 年 12 月、1982 年 1 月和 1982 年 2 月出生的群体设置为 2，1981 年 3 月 - 5 月出生的群体设置为 -1，以此类推。这与刘生龙和胡鞍钢 (2018)、陈技伟和冯帅章 (2022) 的做法一致。1981 年出生的群体在 2008 年为 27 岁，而受高校扩招影响的群体在 2003 年后开始进入劳动力市场。考虑到刚进入劳动力市场的实习期和工作经验不足，同时与断点回归的局部平均处理效应相呼应，本文将出生季度限定在参照点前后 20 个季度，即出生于 1976 年 6 月 - 1986 年 8 月间 (25 ~ 41 岁) 的群体，这一

^① Vytlačil (2002) 证明，模糊断点回归 (RD) 与 IV 的识别条件和识别对象完全一致。

年龄段群体受生命周期收入波动影响也相对较小。

通过以上设定,本文构造是否受高校扩招影响的虚拟变量 *Expan*,当出生季度虚拟变量大于 0(出生于 1981 年 9 月及以后)时取值为 1,表示受到高校扩招影响,小于等于 0(出生于 1981 年 8 月及以前)时取值为 0,表示未受到高校扩招影响。表 1 报告了变量的描述性统计,不同学历劳动者的相对教育回报可能受出生组群因素(Cohort Effect)的影响,即不同出生年龄段的群体,除了受教育程度不同外,还可能因为与出生组群相关的其他因素而存在较大的收入差距,如不同出生年龄段群体进入劳动力市场时的工作机会可能存在系统性差异。Card & Lemieux(2001)建议比较高中与本科学历劳动者的教育回报差异来控制出生组群效应,本科学历劳动者比高中学历劳动者多 4 年(专科则 3 年)受教育年限,但也缺少 4 年(3 年)工作经验,这两类群体测算出的教育回报与实质性的大学回报率是较为接近的。本文除采用 Card & Lemieux(2001)的建议外,还将对照组扩展到初中学历,以考虑不同受教育程度的教育回报异质性。

表 1 变量描述性统计

变量	说明	观测值	均值	标准差	最小值	最大值
<i>lnwage</i>	月收入对数(2011 年不变价)	306564	7.977	0.506	6.416	9.551
<i>Expan</i>	是否受到扩招影响(否=0,是=1)	306564	0.554	0.497	0	1
<i>Exp</i>	期望工作经验(年龄-受教育年限-6)	306564	16.88	4.395	4	27
<i>Ethnics</i>	民族(非汉族=0,汉族=1)	306564	0.939	0.239	0	1
<i>Gender</i>	性别(男性=0,女性=1)	306564	0.398	0.489	0	1
<i>Hukou</i>	户口(农村=0,城市=1)	306564	0.206	0.405	0	1
<i>Marriage</i>	婚姻状况(未婚=0,已婚=1)	306564	0.885	0.320	0	1

资料来源:根据 2011-2017 年中国流动人口动态监测调查(CMDS)数据计算得到。

(三) 流动人口高等教育溢价变动的趋势:2011-2017

中国城镇住户调查(UHS)的样本仅限于城市层面,2009 年前也可能因扩招初期大学生增加的比例不高、需求侧因素强劲、农村非熟练劳动力向城市转移等原因,导致高等教育溢价并未大幅降低。为测算近期高等教育溢价,本文采用了 2011-2017 年 CMDS 数据,样本为城市地区的流动人口。调查对象为流动人口,导致本文的结果可能存在样本选择误偏。但由于本文研究流动人口高等教育溢价,样本范围为 1976 年 6 月-1986 年 8 月出生的群体,工具变量设计要求将受教育程度限制在至少接受过初中教育的群体。从这两个角度而言,这部分群体更可能流动到城市地区,因此缓解了可能存在的样本选择误偏问题。在受教育程度的度量上,部分研究直接采用受教育年限,

可能忽视了不同层级教育的异质性回报。本文通过不同的实验组和对照组选取尝试探究不同层级教育回报异质性。

表2报告了不同年份、不同学历组合下 OLS 和 IV 的估计结果，点估计值和 95% 置信区间分别绘制在图 1 至图 4 中。图 1 刻画了 OLS 和 IV 估计的本科相对于高中教育溢价随时间的变动趋势，教育溢价反映的是经历 4 年本科教育的群体相对于高中学历但多 4 年工作经历群体的教育回报，IV 识别“政策遵从者”的教育溢价。从 IV 估计结果来看，2011 年本科相对于高中的教育溢价为 35.8%，按高中受教育年限为 12 年而本科受教育年限为 16 年来算，本科教育在 2011 年享有年均 8.95% 的教育溢价，这一回报率在 Card (1999) 总结的 6% ~ 10% 区间内。但流动人口高等教育溢价整体呈逐年下滑趋势，在 2011 - 2014 年间急剧下降，除在 2016 年有小幅不显著的回升外，一直降低到 2017 年的 22.7%，平均每年降低约 2.2%，2017 年的年均教育回报仅为 5.68%。对比 OLS 与 IV 的估计结果，二者的点估计值和置信区间差异不大，且呈现类似的逐年降低趋势。

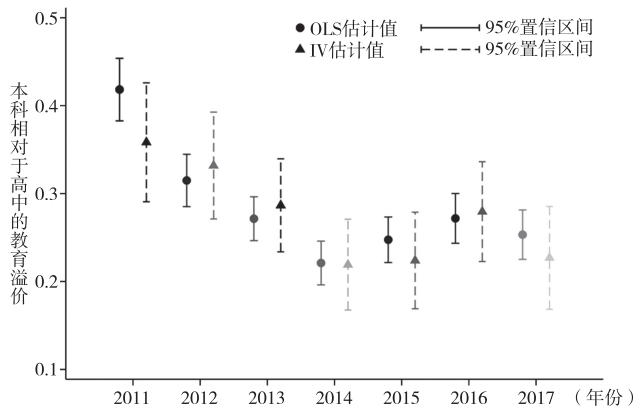


图 1 本科相对于高中的教育溢价

注：该图展示了 OLS 和 IV 估计的不同年份本科相对于高中的教育溢价及其 95% 置信区间。计算年均教育溢价时，本科、专科、高中和初中的受教育年限分别设定为 16 年、15 年、12 年和 9 年，下同。相比于高中，本科在 2011 年和 2017 年的年均教育溢价分别为 8.95% 和 5.68%。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查 (CMDS) 数据计算得到。

图 2 将未上大学群体扩展到初中学历，刻画本科相对于高中或初中的教育溢价，OLS 和 IV 估计结果均呈逐年降低趋势。从数值上对比图 1 与图 2 的 OLS 估计结果，二者差异不大，但图 2 的 IV 估计结果却明显高于图 1，教育回报率测算中 IV 估计结果高于 OLS 估计结果的现象 (Card, 1999) 再次出现，可能有四个原因：一是能力更低的

人倾向上大学（上大学对于能力更高的人机会成本也更高）；二是测量误偏（核心解释变量在测量上的误差产生的衰减误偏导致 OLS 估计结果下偏）；三是发表误偏（研究者或期刊更倾向报告或发表教育回报较高的结果）；四是异质性回报（工具变量对应的“政策遵从者”教育回报更高）。对比图 1 与图 2 的 IV 估计结果，本文为异质性教育回报导致 IV 估计结果高于 OLS 估计结果提供了一些证据：当非高等教育劳动者样本中包含初中学历时，初中学历劳动者中的“政策非遵从者”（Never Takers）比重更高^①，导致 IV 估计出的“政策遵从者”教育溢价高于 OLS 估计的教育溢价。从 IV 估计结果来看，相比于高中或初中，本科在 2011 年和 2017 年的年均教育溢价分别为 11.5% 和 8.6%。

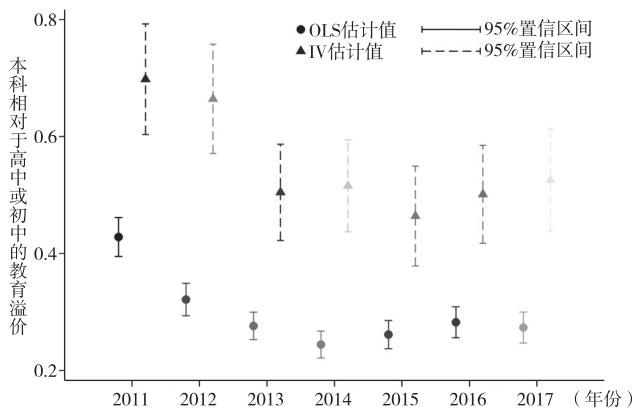


图 2 本科相对于高中或初中的教育溢价

注：通过描述性统计进行加权平均可得，高中或初中样本中，平均受教育年限为 9.9 年（ $9 \times 57.4\% / (57.4\% + 25.09\%) + 12 \times 25.09\% / (57.4\% + 25.09\%)$ ），相比于高中或初中，本科在 2011 年和 2017 年的年均教育溢价分别为 11.5% 和 8.6%。

资料来源：根据 2011—2017 年中国流动人口动态监测调查（CMDS）数据计算得到。

图 3 刻画了本科或专科相对于高中的教育溢价，相比于高中，IV 估计结果表明本科或专科相对于高中的年均教育溢价在 2011 年和 2017 年分别为 8.18% 和 4.65%，OLS 估计结果相对于 IV 估计结果差异不大。对比图 1 和图 3，说明本科相对于专科具有更高的年均教育溢价，这也说明在测算教育回报时使用教育虚拟变量比受教育年限量受教育程度的做法更加准确。最后，在图 4 中（小时收入的结果见图 5），本文报告了本科或专科相对于高中或初中的教育溢价，IV 估计结果表明，相比于高中或初中，本

^① 即无论高校是否扩招均不上大学的群体。

科或专科在 2011 年和 2017 年的年均教育溢价分别为 8.09% 和 5.60%，IV 估计结果仍显著高于 OLS 估计结果，这与工具变量识别“政策遵从者”的解释是一致的。

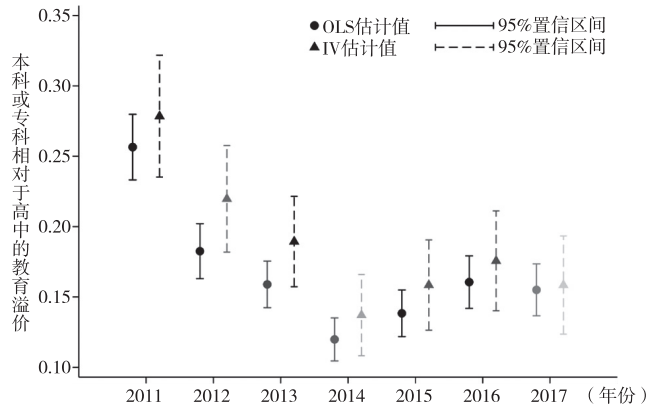


图3 本科或专科相对于高中的教育溢价

注：通过描述性统计进行加权平均可得，本科或专科样本中，平均受教育年限为 15.4 年 ($15 \times 10.66\% / (10.66\% + 6.85\%) + 16 \times 6.85\% / (10.66\% + 6.85\%)$)，相比于高中，本科或专科在 2011 年和 2017 年的年均教育溢价分别为 8.18% 和 4.65%。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查 (CMDS) 数据计算得到。

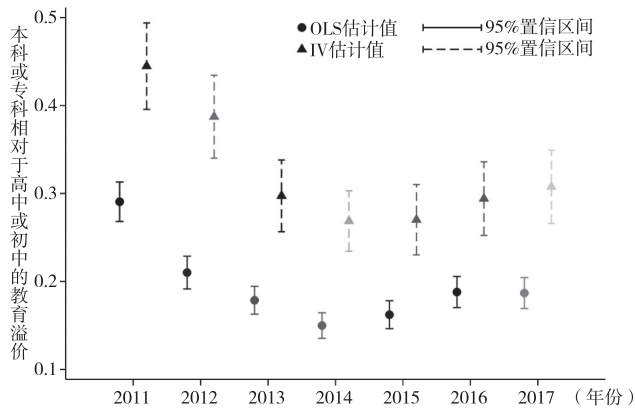


图4 本科或专科相对于高中或初中的教育溢价

注：通过描述性统计进行加权平均可得，高中或初中样本中，平均受教育年限为 9.9 年，本科或专科样本中，平均受教育年限为 15.4 年，相比于高中或初中，本科或专科在 2011 年和 2017 年的年均教育溢价分别为 8.09% 和 5.60%。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查 (CMDS) 数据计算得到。

上述结果可总结为：第一，无论怎样对教育程度进行分类，流动人口高等教育相对于非高等教育的溢价始终降低了，这一降低程度在 2011 - 2014 年间更加明显，可能

的原因在于随着农村劳动力向城市转移速度放缓，劳动力无限供给特征逐渐消失，中国经济在 2004 - 2010 年间已经越过了“刘易斯拐点”（蔡昉，2018），市场对非高等教育劳动者需求的增加高于供给速度。随着中国高等教育的不断扩张，高等教育已经演变成一种大众教育，教育回报下降，最终导致高等教育劳动者相比于非高等教育劳动者的教育溢价逐年减少。第二，在测算教育回报过程中，以受教育年限量受教育程度可能忽略了不同教育层级教育回报的系统性差异。第三，将样本限定在高中及以上学历时，OLS 与 IV 测算的流动人口高等教育溢价差异不大，而将非高等教育扩展到初中时，IV 的测算结果明显高于 OLS 的测算结果，这印证了 IV 识别“政策遵从者”教育回报的结论。

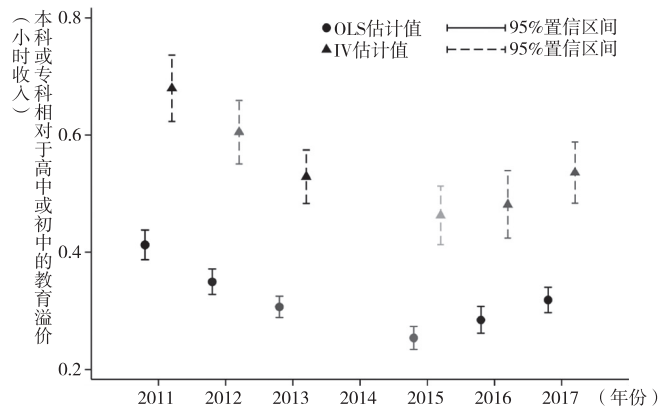


图 5 本科或专科相对于高中或初中的教育溢价（小时收入）

注：图 5 与图 4 一致，仅将被解释变量更换为小时收入，2014 年的工作时间数据缺失。

资料来源：根据 2011 - 2013 年、2015 - 2017 年中国流动人口动态监测调查（CMDS）数据计算得到。

表 2 流动人口高等教育溢价的 OLS 和 IV 估计结果

年份	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2011	2012	2013	2014	2015	2016	2017
A 部分：本科相对于高中							
OLS 估计值	0.418 *** (0.018)	0.315 *** (0.015)	0.271 *** (0.013)	0.221 *** (0.013)	0.247 *** (0.013)	0.272 *** (0.014)	0.253 *** (0.014)
2SLS 估计值	0.358 *** (0.035)	0.332 *** (0.031)	0.287 *** (0.027)	0.219 *** (0.026)	0.224 *** (0.028)	0.279 *** (0.029)	0.227 *** (0.030)

续表

年份	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	2011	2012	2013	2014	2015	2016	2017
第一阶段 KP F 值	457.4	774.9	992.5	1284.9	1120.4	1213.8	1116.9
控制变量	是	是	是	是	是	是	是
观测值数	6463	13563	17536	17205	16312	13941	12896
调整后 R ²	0.329	0.185	0.190	0.190	0.192	0.194	0.210
B 部分：本科相对于高中或初中							
OLS 估计值	0.428 *** (0.017)	0.321 *** (0.014)	0.276 *** (0.012)	0.244 *** (0.012)	0.262 *** (0.012)	0.283 *** (0.013)	0.274 *** (0.013)
2SLS 估计值	0.698 *** (0.048)	0.664 *** (0.048)	0.505 *** (0.042)	0.516 *** (0.040)	0.464 *** (0.044)	0.501 *** (0.043)	0.525 *** (0.045)
第一阶段 KP F 值	344.4	544.1	685.8	1047.3	848.9	1126.4	1070.0
控制变量	是	是	是	是	是	是	是
观测值数	16806	38370	50489	50857	46611	37034	33705
调整后 R ²	0.264	0.118	0.148	0.139	0.145	0.142	0.168
C 部分：本科或专科相对于高中							
OLS 估计值	0.257 *** (0.012)	0.183 *** (0.010)	0.159 *** (0.008)	0.120 *** (0.008)	0.138 *** (0.008)	0.161 *** (0.010)	0.155 *** (0.009)
2SLS 估计值	0.278 *** (0.022)	0.220 *** (0.019)	0.189 *** (0.016)	0.137 *** (0.015)	0.158 *** (0.016)	0.176 *** (0.018)	0.158 *** (0.018)
第一阶段 KP F 值	963.5	1728.3	2651.7	2651.7	2273.1	2130.9	2081.4
控制变量	是	是	是	是	是	是	是
观测值数	8752	17642	22761	23564	21518	18743	17628
调整后 R ²	0.300	0.178	0.184	0.183	0.185	0.181	0.201
D 部分：本科或专科相对于高中或初中							
OLS 估计值	0.291 *** (0.011)	0.210 *** (0.010)	0.179 *** (0.008)	0.150 *** (0.007)	0.162 *** (0.008)	0.188 *** (0.009)	0.187 *** (0.009)
2SLS 估计值	0.445 *** (0.025)	0.387 *** (0.024)	0.297 *** (0.021)	0.269 *** (0.018)	0.270 *** (0.020)	0.294 *** (0.021)	0.308 *** (0.021)
第一阶段 KP F 值	1020.8	1718.3	2211.6	2930.6	2422.0	2413.5	2436.4
控制变量	是	是	是	是	是	是	是
观测值数	19095	42449	55714	57216	51817	41836	38437
调整后 R ²	0.267	0.130	0.154	0.150	0.152	0.150	0.179

注：括号中为稳健标准误；* $p < 0.1$ ，** $p < 0.05$ ，*** $p < 0.01$ ；控制变量包含性别、期望工作经验及其平方项、民族、婚姻状况、户口和省份固定效应。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查 (CMDS) 数据计算得到。

四 流动人口高等教育溢价变动的解释

(一) 流动人口高等教育相对供给与教育溢价：特征事实

在发现流动人口高等教育溢价降低的基础上，本文进一步解释流动人口高等教育溢价变动的原因，研究对象集中在流动人口。图6显示流动人口高等教育相对供给整体呈增长趋势^①，流动人口高等教育相对供给与高等教育溢价之间呈现“剪刀”型特征。由此，本文尝试探究流动人口高等教育相对供给变动对高等教育溢价的影响。

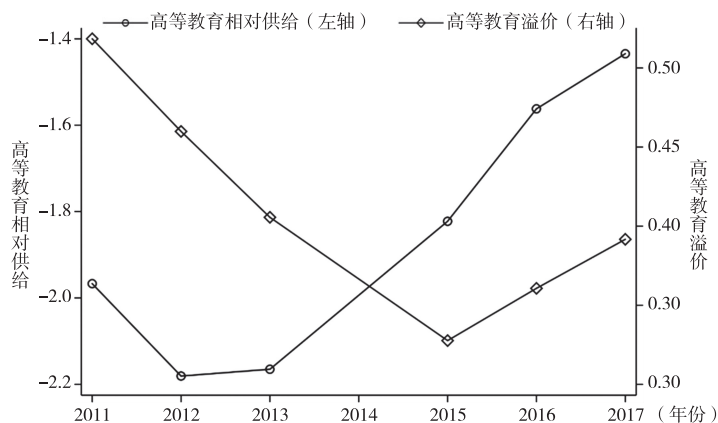


图6 流动人口高等教育相对供给与溢价变动

注：流动人口高等教育相对供给是指本科或专科学历劳动者与高中或初中学历劳动者总工作小时数的比值（对数），权重为全样本小时收入差异：专科学历劳动者1小时劳动供给相当于本科学历劳动者0.88小时劳动供给，初中学历劳动者1小时劳动供给相当于高中学历劳动者0.77小时劳动供给；流动人口高等教育溢价是指本科或专科学历劳动者与高中或初中学历劳动者的小时收入比值（对数）。

资料来源：根据2011-2013年、2015-2017年中国流动人口动态监测调查（CMDS）数据计算得到。

将研究对象集中在流动人口，需要考虑流动决策与教育溢价之间的反向因果关系和遗漏变量导致的内生性问题，即高等教育劳动者倾向于流向高等教育溢价更高的地区，反向因果关系导致出现高等教育劳动者相对供给越高的地区，高等教育溢价反而更高的“伪回归”现象。图7刻画了流动人口高等教育溢价与高等教育相对供给之间

^① 2011年的流动人口高等教育相对供给略高于2012年的可能原因在于，2011年的CMDS数据样本量较其他年份偏少，在21世纪以来高校不断扩招的事实下，有理由判定流动人口高等教育相对供给逐年增加。

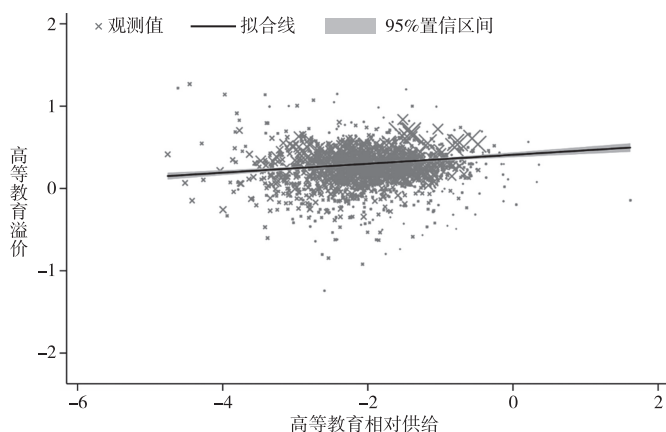


图7 流动人口高等教育溢价与相对供给的关系（城市层面不区分年龄段）

注：横轴为流动人口高等教育相对供给，纵轴为流动人口高等教育溢价，具体测算过程见正文，权重为总工作小时数，直线为OLS拟合线；图8除区分年龄段 j 外，其余类似。

资料来源：根据2011 - 2013年、2015 - 2017年中国流动人口动态监测调查（CMDS）数据计算得到。

的关系，可以发现流动人口高等教育溢价与高等教育相对供给呈正相关关系，即要素相对供给的提升反而提升了该要素的回报，这一现象容易被解读为高等教育的溢出效应。为消除图7带来的困惑，图8将样本按城市—年龄段—年份进行划分，在城市一年份内引入不同年龄段后，流动人口高等教育溢价与相对供给呈现负相关关系。不过，从简单的OLS估计值来看，流动人口高等教育相对供给提升10%，高等教育溢价降低约0.7%，由此测算出的流动人口高等教育与非高等教育的替代弹性高达14，远高于文献上一致认可的1.4（Katz & Murphy, 1992；Goldin & Katz, 2010）。下文将构建一个包含不同年龄段、不同受教育程度劳动者不完全替代模型，并构造工具变量缓解流动人口高等教育相对供给的内生性，尝试解释流动人口高等教育溢价的变动。

（二）理论模型与实证设计

本文采用与Katz & Murphy（1992）、Card & Lemieux（2001）和Manacorda et al.（2012）类似的框架展开分析。由于重点关注不同年龄段之间、流动人口与本地人口之间的不完全替代性，本文参照马光荣等（2017）的研究，将生产函数写为如下嵌套的CES函数形式：

$$y_t = A_t \left[\sum_j \alpha_j L_{jt}^\eta \right]^{1/\eta} \quad (3)$$

其中 j 和 t 分别表示年龄段和年份， y_t 为产出， A_t 为中性技术进步， L_{jt} 表示 j 年龄

段 t 年份的劳动供给^①, α_j 是 L_j 的生产效率, 不同年龄段劳动者之间的替代弹性 $\sigma_A = 1/(1 - \eta)$ 。

L_j 设定为如下 CES 函数形式:

$$L_j = [\theta_{sjt} S_{jt}^\rho + \theta_{ujt} U_{jt}^\rho]^{1/\rho} \quad (4)$$

其中 S_{jt} 和 U_{jt} 分别表示 j 年龄段 t 年份的高等教育 (含本科专科) 和非高等教育 (含高中初中) 劳动者供给, θ_{sjt} 和 θ_{ujt} 分别对应两类劳动的生产效率, 不同受教育程度劳动者间的替代弹性 $\sigma_E = 1/(1 - \rho)$, σ_E 越大, 不同受教育程度劳动者之间越容易替代, σ_E 无穷大时二者完全替代。将式 (4) 带入式 (3), 令不同劳动者的收入等于其边际产出, 高等教育劳动者收入 (w_{sjt}) 与非高等教育劳动者收入 (w_{ujt}) 分别为:

$$w_{sjt} = \frac{dy_t}{dS_{jt}} = A_t \alpha_j L_j^{\eta-\rho} \Psi_t \theta_{sjt} S_{jt}^{\rho-1} \quad (5)$$

$$w_{ujt} = \frac{dy_t}{dU_{jt}} = A_t \alpha_j L_j^{\eta-\rho} \Psi_t \theta_{ujt} U_{jt}^{\rho-1} \quad (6)$$

其中, $\Psi_t = [\sum_j \alpha_j L_j^\eta]^{1/\eta-1}$, 联立式 (5) 和 (6), j 年龄段 t 年份的高等教育和非高等教育劳动者相对收入 $\left(\ln w_{sjt} = \ln\left(\frac{w_{sjt}}{w_{ujt}}\right)\right)$ 与相对供给 $\left(\ln s_{jt} = \ln\left(\frac{S_{jt}}{U_{jt}}\right)\right)$ 和技术进步 $\left(\ln\left(\frac{\theta_{sjt}}{\theta_{ujt}}\right)\right)$ 的关系为:

$$\ln\left(\frac{w_{sjt}}{w_{ujt}}\right) = \ln\left(\frac{\theta_{sjt}}{\theta_{ujt}}\right) - \frac{1}{\sigma_E} \ln\left(\frac{S_{jt}}{U_{jt}}\right) \quad (7)$$

式 (7) 即为经典的技术进步与教育供给影响工资差距的理论基础: 偏向高技能劳动者的技术进步提升高低技能劳动者的收入差距, 高技能劳动者的相对供给降低高低技能劳动者的收入差距 (Goldin & Katz, 2010)。本文使用城市层面的数据, 式 (7) 对应的实证模型为:

$$\ln\left(\frac{w_{scjt}}{w_{ucjt}}\right) = \frac{-1}{\sigma_E} \ln\left(\frac{S_{cjt}}{U_{cjt}}\right) + b_{pt} + \epsilon_{cjt} \quad (8)$$

式 (8) 相对于式 (7) 增加了城市 c 标记, 技术进步项隐含在固定效应 b_{pt} 中, 包括省份、年份、省份 \times 年份固定效应, ϵ_{cjt} 为扰动项。

(三) 变量的构造与工具变量设定

本文将 26 ~ 59 岁的流动人口每 5 岁一段分成 7 个年龄段 (最后一个年龄段为 56 ~

^① 本文的研究对象为流动人口, 若考虑本地人口, 在生产函数基础上再嵌套一层流动人口和本地人口的 CES 生产函数, 本地人口在求导、取比值的过程中被消除了, 结论不变。

59岁)。高等教育包含本科与专科,非高等教育包含高中与初中,本科与专科、高中与初中劳动者的相对劳动供给效率存在差异,本文参照Katz & Murphy (1992)、Card & Lemieux (2001)的做法,通过2011-2017年本科与专科学历劳动者小时收入差异作为权重。通过计算,专科学历劳动者1小时劳动供给相当于本科学历劳动者0.88小时劳动供给,而初中学历劳动者1小时劳动供给相当于高中学历劳动者0.77小时劳动供给。本文在城市一年龄段一年份子样本中加权累加月工作小时数分别构造 S_{cjt} 和 U_{cjt} ,类似地,通过子样本中的平均小时收入得到 w_{scjt} 和 w_{ucjt} 。不同学历劳动者相对收入对数($\ln w_{cjt}$)和相对供给($\ln s_{cjt}$)的关系如图8所示,高等教育相对非高等教育的教育溢价与相对供给负相关,由于(非)高等教育劳动者更倾向于流向(非)高等教育溢价更高的地区($cov\left(\ln\left(\frac{S_{cjt}}{U_{cjt}}\right), \epsilon_{cjt}\right) > 0$)。图8显示的负相关关系(绝对值)可能存在低估,真实情况下流动人口高等教育相对供给增加导致的高等教育溢价降低程度可能更加明显。以上逻辑通过对式(8)进行一些简单推导即得式(9)。

$$-\left(\frac{1}{\sigma_E}\right)_{ols} = -\left(\frac{1}{\sigma_E}\right)_{真实} + cov\left(\ln\left(\frac{S_{cjt}}{U_{cjt}}\right), \epsilon_{cjt}\right) \quad (9)$$

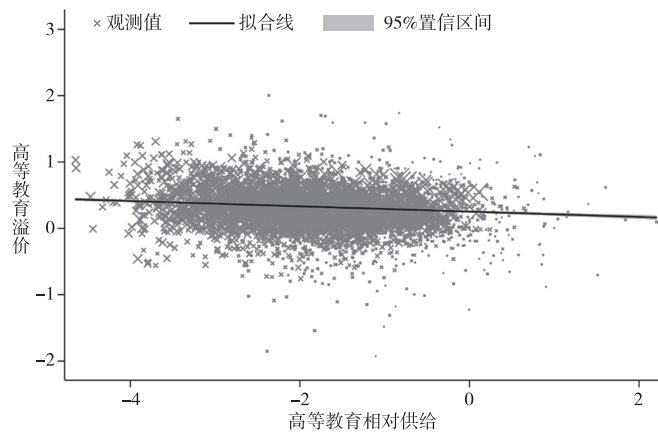


图8 流动人口高等教育溢价和相对供给的关系(城市层面区分年龄段)

资料来源:根据2011-2013年、2015-2017年中国流动人口动态监测调查(CMDS)数据计算得到。

$-\frac{1}{\sigma_E}$ 的OLS估计值 $\left(-\left(\frac{1}{\sigma_E}\right)_{ols}\right)$ 为其真实值 $\left(-\left(\frac{1}{\sigma_E}\right)_{真实}\right)$ 与 $cov\left(\ln\left(\frac{S_{cjt}}{U_{cjt}}\right), \epsilon_{cjt}\right)$ 之和,

由于 $cov\left(\ln\left(\frac{S_{cjt}}{U_{cjt}}\right), \epsilon_{cjt}\right) > 0$, $\left(\frac{1}{\sigma_E}\right)_{ols}$ 存在低估。为缓解遗漏变量导致的内生性问题, 本文借鉴 Card (2001, 2009) 的方法预测不同城市流动人口高等教育相对供给 ($\ln sr_{cjt}$), 将预测值作为真实值的工具变量。核心逻辑是新一代流动人口由于关系网络等原因, 更倾向于流向同一流出地流动人口所在的城市或地区。构造 $\ln sr_{cjt}$ 预测值可分为五步: 第一, 在 CMDS 数据中统计 M_{mcjt} : m 省份 t 年流向 c 城市的 j 年龄段初中及以上学历 (不含研究生, 下同) 的劳动者数量; 第二步, 在 2005 年 1% 人口抽样调查微观数据中计算 $\lambda_{mc,2005}$: 2005 年 m 省份流向 c 城市的流动人口中本科与专科学历劳动者占初中及以上学历劳动者的比重; 第三步, 计算 $\lambda_{mc,2005}$ 与 M_{mcjt} 的乘积, 作为 m 省份 t 年流向 c 城市的 j 年龄段本科与专科学历劳动者数量预测值, 背后的逻辑是流入地和流出地相同的流动人口, 学历结构之间存在相关性, 该步骤不要求流动人口数量外生; 第四步, 对不同流出省份 (m) 的本科与专科学历劳动者预测值累加, 得到不同城市 (c)、不同年龄段 (j) 和不同年份 (t) 的本科与专科学历劳动者 S_{cjt} 预测值; 第五步, 对高中与初中学历劳动者, 重复前四个步骤, 即可得到不同城市 (c)、不同年龄段 (j) 和不同年份 (t) 的高中与初中学历劳动者 U_{cjt} 预测值, 取二者比值的对数即为 $\ln sr_{cjt}$ 预测值 $\left(\ln \frac{S_{cjt} \text{ 预测值}}{U_{cjt} \text{ 预测值}}\right)$ 。

$$S_{cjt} \text{ 预测值} = \left(\sum_m \lambda_{mc,2005} M_{mcjt}\right), m = 1, 2, \dots, 31 \quad (10)$$

$$U_{cjt} \text{ 预测值} = \left(\sum_m (1 - \lambda_{mc,2005}) M_{mcjt}\right), m = 1, 2, \dots, 31 \quad (11)$$

通过上述步骤, $\ln sr_{cjt}$ 预测值与真实值的关系如图 9 所示, 二者存在正相关关系。本文使用 $\ln sr_{cjt}$ 预测值作为 $\ln sr_{cjt}$ 真实值的工具变量。作为交叉验证, 本文还考虑了不区分年龄段的情形, 即基于流动“惯性”构造 $\ln sr_{ct}$ 的预测值, 结果见图 10, 此时同一预测值对应的真实值也更集中。因此, 工具变量的“第一阶段”也更加明显, 后文同时汇报了两类情形的估计结果, 二者十分接近。

(四) 结果与分析

对式 (8) 的 OLS 和 IV 估计结果如表 3 所示, 第 (1) ~ (4) 列为相同教育不同年龄段完全替代情形下的估计结果。第 (1) 列的结果对应图 7, 流动人口高等教育相对供给每提升 10%, 高等教育溢价反而提升 0.54%, 这一正相关关系是明显的“伪回归”现象, 反映了遗漏变量导致的内生性问题。第 (2) 列同时控制了年份、省份、年份 × 省份固定效应, OLS 估计的系数差异不大。第 (3) 列、第 (4) 列使用 $\ln sr_{ct}$ 预测值作为真

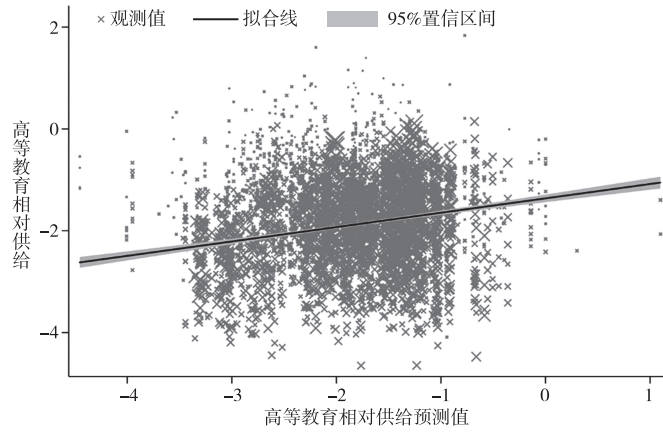


图9 高等教育相对供给预测值与真实值的关系（城市层面区分年龄段）

注：横轴为流动人口高等教育相对供给预测值，纵轴为其真实值。

资料来源：根据2011-2013年、2015-2017年中国流动人口动态监测调查（CMDS）数据和2005年1%人口抽样调查微观数据计算得到。

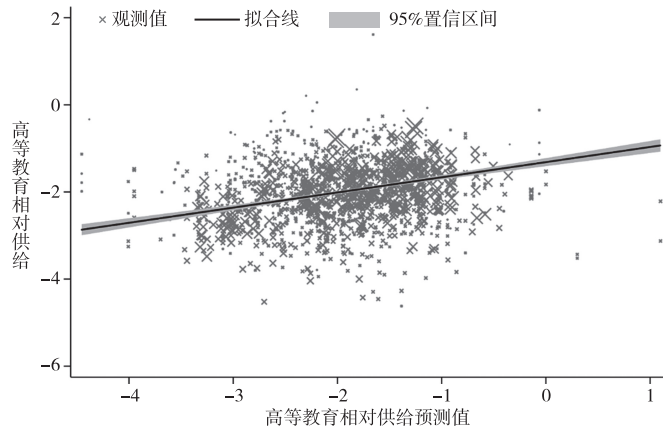


图10 高等教育相对供给预测值与真实值的关系（城市层面不区分年龄段）

注：横轴为流动人口高等教育相对供给预测值，纵轴为其真实值。

资料来源：根据2011-2013年、2015-2017年中国流动人口动态监测调查（CMDS）数据和2005年1%人口抽样调查微观数据计算得到。

实值的工具变量，第（3）列第一阶段估计结果对应图10， $\ln sr_{ct}$ 预测值每提升10%， $\ln sr_{ct}$ 提升3.48%，第一阶段F值远大于经验值10，说明不存在弱工具变量问题，2SLS估计值相对于OLS估计值的符号发生反转，流动人口高等教育相对供给每提升10%，高等教育溢价降低1.12%，OLS估计值可能因为流动人口的选择性迁移而向上

误偏。

第 (5) ~ (8) 列为相同教育不同年龄段不完全替代情形下的估计结果, 第 (5) 列的结果对应图 8, 流动人口高等教育相对供给每提升 10%, 高等教育溢价降低 0.4%, 尽管价格与供给呈现符合逻辑的负相关关系, 但在数值上仍可能因为遗漏变量问题而向上误偏, 第 (6) 列控制年份、省份、省份 \times 年份固定效应后, OLS 估计系数由 -0.04 降低为 -0.07, 与遗漏变量的解释相符。第 (7) 列、第 (8) 列使用 $\ln sr_{cjt}$ 预测值作为真实值的工具变量, 第 (7) 列第一阶段估计结果对应图 9, 第一阶段估计结果与相同教育不同年龄段完全替代情形差异不大, 2SLS 估计结果显示流动人口高等教育相对供给每提升 10%, 高等教育溢价降低约 1.3%, 相同教育不同年龄段是否完全替代对结果影响不大。

对比现有文献, Ottaviano & Peri (2012) 总结美国的 $1/\sigma_E$ 的估计值在 0.3 和 0.7 之间, 本文估计的流动人口高等教育与非高等教育劳动者的替代弹性约为 7 ($1/\sigma_E = 0.13$), 高于现有文献和马光荣等 (2017) 基于 2009 年以前 UHS 数据的估计值, 这一差异产生的一个重要原因在于本文将专科归为高等教育, 而不是按一定比例将专科学历劳动供给分割到高等与非高等教育。本文计算的全国层面 25 ~ 41 岁年龄段流动人口高等教育相对供给 $\ln sr_{25-41,t}$ 从 2011 年的 -1.69 增加到 2017 年的 -1.10, 由供给导致的高等教育溢价下降 7.67% ($-(1.69 - 1.10) \times 0.13$), 而同一时期 IV 估计的同一年龄段高等教育溢价 (小时收入) 由 0.680 降低到 0.536, 降低了 14.4%, 流动人口高等教育的供给可解释这一时期高等教育溢价降低的 53.2% ($7.67\% / 14.4\%$)。

表 3 的基准结果同时考虑了男性和女性劳动者, 按照惯例 (Katz & Murphy, 1992; Card & Lemieux, 2001)^①, 本文仅使用男性样本重复了上述步骤, 结果见表 4。男性劳动者的结果与全样本结果一致, 流动人口高等教育和非高等教育男性劳动者的替代弹性约为 5, 低于全样本的估计结果。替代弹性越小, 意味着不同劳动者间可替代程度更低, 相对于全样本而言, 男性高等教育与非高等教育劳动者间的可替代性更低。

为保证结果的稳健性, 本文进一步引入了一系列城市层面控制变量, 包括人均 GDP、城市人口规模、城市登记失业率、第二产业就业比重、第三产业就业比重和地方政府财政支出与收入之比, 结论仍稳健, 估计结果限于篇幅略去。值得一提的是, 引

^① 原因在于女性收入数据本身存在更大的误差, 同时, 女性可能因为生育等原因而出现职业中断, 导致按年龄或工作经验对样本进行划分出现较大偏误。

入城市层面控制变量后，流动人口高等教育相对供给与溢价之间的“伪”正相关现象消失了，工具变量估计结果与基准结果几乎一致。

表 3 流动人口高等教育劳动者相对供给对教育溢价影响的估计（全样本）

年龄段设定	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	相同教育不同年龄段完全替代				相同教育不同年龄段不完全替代			
方法	OLS	OLS	IV	IV	OLS	OLS	IV	IV
$\ln sr_{ci}(\ln sr_{cj})$	0.054 *** (0.014)	0.057 *** (0.012)	-0.112 *** (0.034)	-0.113 *** (0.035)	-0.040 *** (0.008)	-0.070 *** (0.006)	-0.129 *** (0.029)	-0.132 *** (0.031)
年份##省份		是				是		
年份				是				是
第一阶段回归结果			0.348 *** (0.037)	0.338 *** (0.032)			0.282 *** (0.030)	0.271 *** (0.028)
KP F 值			89.3	111.8			89.4	95.4
观测值数	1539	1539	1444	1444	4479	4479	4292	4292
调整后 R ²	0.034	0.373	0.632	0.651	0.019	0.337	0.598	0.609

注：工具变量构造见正文；年份##省份表示年份、省份、年份×省份固定效应；权重为不同分组中初中及以上学历劳动者的总工作小时数；括号中为稳健标准误；* $p < 0.1$ ，** $p < 0.05$ ，*** $p < 0.01$ 。

资料来源：根据 2011 - 2013 年、2015 - 2017 年中国流动人口动态监测调查（CMDS）数据和 2005 年 1% 人口抽样调查微观数据计算得到。

基准结果以城市划分地理单元，城市间流动人口高等教育溢价差产生的套利空间可能引导流动人口流动，最终使不同地区流动人口高等教育溢价趋向均衡。因此，使用更小层级的地理单元作为研究对象可能低估供给对教育溢价的影响（Borjas, 2003）。为探究地理单元对研究结果的影响，本文将研究地理单元转向省级层面，流动人口高等教育溢价与相对供给的关系如图 11 所示，相比于图 8 地理单元为城市层面的结果，地理单元处于省级层面时，流动人口高等教育相对供给对高等教育溢价的影响（斜率绝对值）更大，流动人口高等教育相对供给每提升 10%，高等教育溢价降低约 3%^①，流动人口高等教育与非高等教育劳动者的替代弹性约为 3.3，测算值在合理范围内，接近主流文献测算区间下限的原因在于，本文将专科学历劳动者归为高等教育劳动者，导致高等教育与非高等教育劳动者之间的替代弹性提升了。使用省份作为地理单元测

① 估计结果限于篇幅未报告。

算的高等教育与非高等教育劳动者替代弹性低于使用城市层面的测算结果，说明基准结果估计的是流动人口高等教育相对供给对溢价影响的上界。

表 4 流动人口高等教育劳动者相对供给对教育溢价影响的估计（男性样本）

年龄段设定	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	相同教育不同年龄段完全替代				相同教育不同年龄段不完全替代			
方法	OLS	OLS	IV	IV	OLS	OLS	IV	IV
$\ln sr_{ijt}$	0.054 *** (0.015)	0.022 * (0.012)	-0.098 *** (0.034)	-0.099 *** (0.035)	-0.031 *** (0.008)	-0.056 *** (0.007)	-0.185 *** (0.044)	-0.190 *** (0.046)
年份##省份		是				是		
年份				是				是
第一阶段回归结果			0.335 *** (0.037)	0.326 *** (0.032)			0.232 *** (0.031)	0.222 *** (0.029)
KP F 值			83.6	103.0			55.2	56.6
观测值数	1369	1369	1245	1245	3413	3413	3246	3246
调整后 R ²	0.030	0.514	0.654	0.671	0.010	0.314	0.514	0.526

注：括号中为稳健标准误；* $p < 0.1$ ，** $p < 0.05$ ，*** $p < 0.01$ 。

资料来源：根据 2011 - 2013 年、2015 - 2017 年中国流动人口动态监测调查（CMDS）数据和 2005 年 1% 人口抽样调查微观数据计算得到。

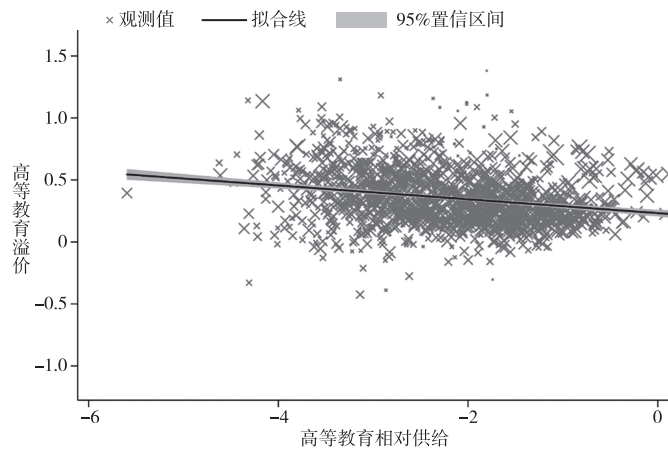


图 11 流动人口高等教育溢价和相对供给的关系（省级层面区分年龄段）

资料来源：根据 2011 - 2013 年、2015 - 2017 年中国流动人口动态监测调查（CMDS）数据计算得到。

(五) 溢出效应

上文从供给角度解释了流动人口高等教育溢价变动，通过构造工具变量对图 8 显示的负相关关系进行了修正。对于工具变量估计结果与 OLS 估计结果之间的差异，本文的实证结果支持遗漏变量叙事：不可观测的遗漏变量同时影响了流动人口高等教育相对供给和溢价，掩盖了流动人口高等教育相对供给对溢价的影响。以具有生产和消费二重属性的基础设施为例，更好的基础设施既吸引了高学历劳动者，同时也可能提升高学历劳动者相对于低学历劳动者的生产力，宜人的气候、方言距离也具有类似的逻辑。对流动人口而言，除遗漏变量外，流动人口高等教育的溢出效应也可能导致 OLS 与 IV 估计结果之间的差异，即流动人口高等教育供给密集的地区，溢出效应 (Acemoglu & Angrist, 2000) 缓解了流动人口高等教育劳动力相对供给对高等教育溢价的负面冲击，本文尝试讨论这一替代性解释。

高等教育溢出效应的产生可分为货币和非货币两类途径 (Acemoglu & Angrist, 2000)，前者的核心思想是更大的市场避免了“敲竹杠”问题，后者强调知识信息交流和干中学等非价格机制。本文用城市层面不同年份的流动人口高等教育相对供给 ($\ln sr_{ct}$) 和人口规模 ($\ln popu_{ct}$) 作为城市规模 (Scale) 的代理变量，在 OLS 模型中引入规模变量及其与高等教育 (S) 的交互项 ($S \times Scale$)。表 5 的结果显示规模变量 (Scale) 本身与收入正相关，且规模变量与高等教育的交互项显著为正，规模变量 $\ln sr_{ct}$ 和 $\ln popu_{ct}$ 的标准误分别为 0.692 和 1.217。以第 (3) 列、第 (4) 列为例，其余分组的结果差别不大， $\ln sr_{ct}$ (Mean = -1.952, S. D. = 0.692) 每提升一个标准误，本科或专科相对于高中学历的溢价提升 3.74% (0.692×0.054)，相当于不考虑规模变量时流动人口高等教育溢价的 23.6% ($3.74\% / 15.84\%$)； $\ln popu_{ct}$ (Mean = 5.484, S. D. = 1.217) 每提升一个标准误，本科或专科相对于高中学历的溢价提升 7.30% (1.217×0.06)，相当于不考虑规模变量时流动人口高等教育溢价的 46.1% ($7.30 / 15.84\%$)，但简单引入交互项可能忽略了其他混杂因素。

本文分别按流动人口高等教育相对供给 ($\ln sr_{ct}$) 和每年城市人口规模将城市十等分，重新估计式 (1) 和 (2)，个人所在城市流动人口高等教育相对供给越高 (或城市人口规模越大)，越可能通过溢出效应提升流动人口高等教育溢价。不同高等教育相对供给 (城市人口规模) 十分位数下的高等教育溢价如图 12 和图 13 所示，在后 80% 分位数内，流动人口高等教育溢价差异不大，而在流动人口高等教育相对供给前 20% 城市，高等教育溢价存在明显的提升，这可能预示着溢出效应仅存在于流动人口高等教育相对供给和城市人口规模前 20% 城市。不过，如图 14 和图 15 所示，高等教育劳

动者在这些城市的相对住房租金也更高。最后，本文将被解释变量更换为收入与支出比值的对数，估计结果图 16 和图 17 所示，可以认为高等教育和非高等教育劳动者的收支比在不同流动人口高等教育相对供给（城市人口规模）的城市不存在显著差异。

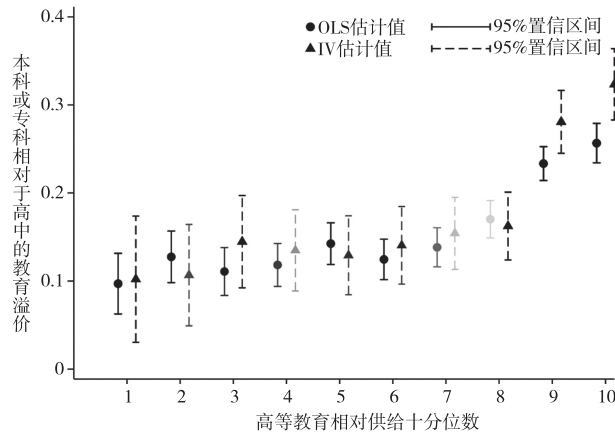


图 12 不同高等教育相对供给十分位数下的流动人口高等教育溢价

注：该图为按流动人口高等教育相对供给将城市十等分后本科或专科相对于高中教育溢价的 OLS 和 IV 估计结果。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查（CMDS）数据计算得到。

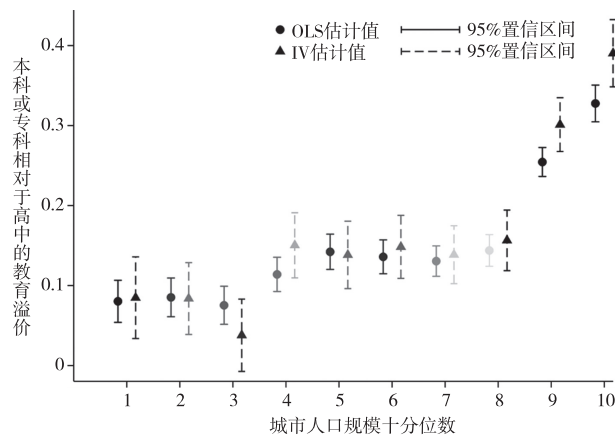


图 13 不同城市人口规模十分位数下的流动人口高等教育溢价

注：该图为按城市人口规模将城市十等分后本科或专科相对于高中教育溢价的 OLS 和 IV 估计结果。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查（CMDS）数据和历年《中国城市统计年鉴》数据计算得到。

表 5 不同城市规模对流动人口高等教育溢价的影响

样本	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	本科相对于高中		本科或专科相对于高中		本科或专科相对于高中或初中		本科相对于高中或初中	
规模变量	相对供给	人口规模	相对供给	人口规模	相对供给	人口规模	相对供给	人口规模
被解释变量	对数收入	对数收入	对数收入	对数收入	对数收入	对数收入	对数收入	对数收入
高等教育	0.365 *** (0.012)	0.168 *** (0.018)	0.256 *** (0.009)	-0.190 *** (0.013)	0.322 *** (0.008)	-0.276 *** (0.011)	0.436 *** (0.011)	-0.283 *** (0.017)
城市规模	0.044 *** (0.004)	0.029 *** (0.002)	0.040 *** (0.004)	0.024 *** (0.002)	0.040 *** (0.002)	0.019 *** (0.001)	0.044 *** (0.002)	0.023 *** (0.001)
高等教育 × 城市规模	0.060 *** (0.007)	0.073 *** (0.003)	0.054 *** (0.005)	0.060 *** (0.002)	0.081 *** (0.004)	0.079 *** (0.002)	0.098 *** (0.006)	0.093 *** (0.003)
控制变量	是	是	是	是	是	是	是	是
观测值数	79081	93276	105078	124493	242455	289525	216458	258308
调整后 R ²	0.235	0.238	0.226	0.228	0.195	0.197	0.190	0.191

注：使用流动人口高等教育相对供给和城市人口规模作为规模变量，在式（2）中引入规模变量及其与流动人口高等教育虚拟变量的交互项；括号中为稳健标准误；* $p < 0.1$ ，** $p < 0.05$ ，*** $p < 0.01$ ；控制变量包含性别、期望工作经验及其平方项、民族、婚姻状况、户口、省份和年份固定效应。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查（CMDS）数据和历年《中国城市统计年鉴》数据计算得到。

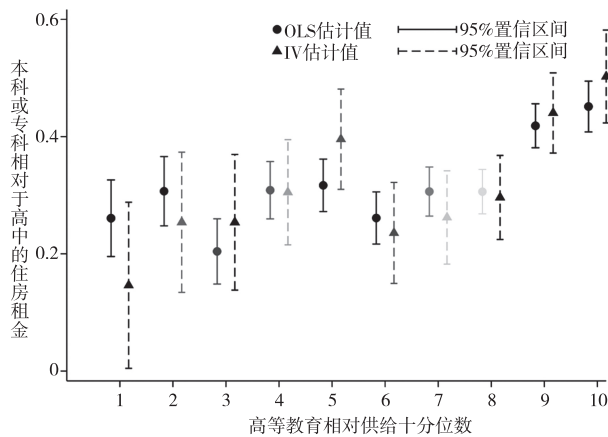


图 14 不同高等教育相对供给十分位数下的住房租金

注：本图为图 12 的被解释变量替换为住房租金（对数）的结果。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查（CMDS）数据计算得到。

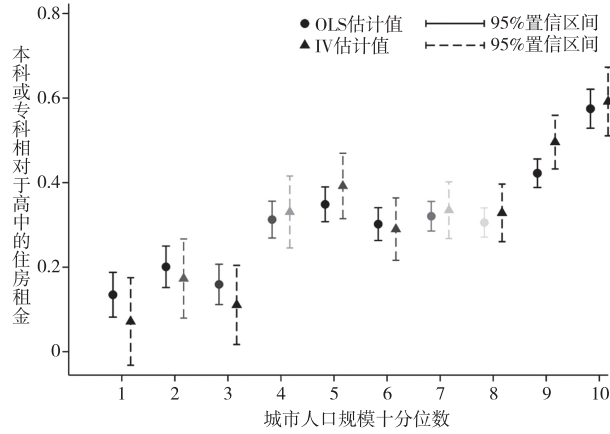


图 15 不同城市人口规模十分位数下的住房租金

注：本图为图 13 的被解释变量替换为住房租金（对数）的结果。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查（CMDS）数据和历年《中国城市统计年鉴》数据计算得到。

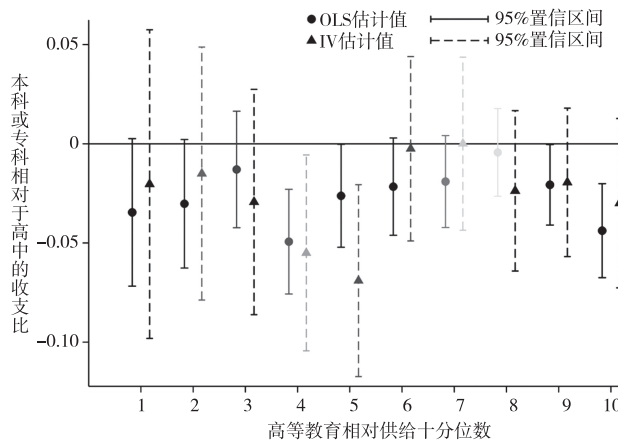


图 16 不同高等教育相对供给十分位数下的收支比

注：本图为图 12 的被解释变量替换为收入支出比值（对数）的结果。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查（CMDS）数据计算得到。

综上，流动人口高等教育可能的溢出效应仅存在于流动人口高等教育相对供给（或城市人口规模）前 20% 城市中，溢出效应可能是由其他因素引起的。即使溢出效应存在，也可能被伴随的生活成本抵消。

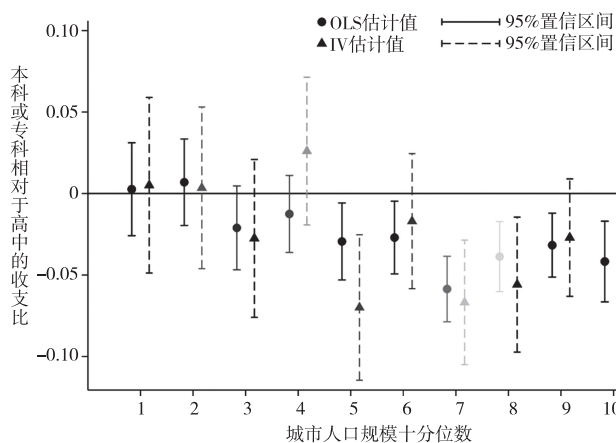


图 17 不同城市人口规模十分位数下的收支比

注：本图为图 13 的被解释变量替换为收入支出比值（对数）的结果。

资料来源：根据 2011 - 2017 年中国流动人口动态监测调查（CMDS）数据和历年《中国城市统计年鉴》数据计算得到。

五 结论与启示

21 世纪以来中国的高等教育经历了持续扩招，高等教育溢价在经历一段上升期后，出现下降趋势。供需规律为此提供了较好的解释，特定年龄段流动人口高等教育相对供给增加 10%，该年龄段流动人口高等教育溢价降低约 1.3% ~ 3%。特大城市即使存在工资溢出效应，也被更高的住房租金等支出抵消，不同规模城市高等教育与非高等教育劳动者的收支比趋于均衡。

高等教育溢价的显著降低主要是由高等教育劳动者供给不断增加导致的，高等教育扩招应考虑教育溢价降低带来的成本及背后的再分配效应。提升高等教育溢价的路径在于继续促进经济结构转型升级，提升对高学历劳动者的需求。不过，中国大学录取率在 2018 年已接近 80%，接受过高等教育的大学毕业生已成为“重点就业群体”，大学是否还是劳动力市场上的稀缺资源，仍然有待讨论。高等教育溢价的降低，也可能源于大学师资和经费的投入增加滞后于大学规模的扩张，从而导致大学教育质量降低。对于这一情形，应根据扩招规模，在合理的范围内加大教育财政资金投入，提升教育经费的使用效率。高等教育溢价的降低，还可能是大学教育与市场需求之间的错配导致的，由于对大学教育的需求和定价已逐步由市场来完成，若大学在专业设置和

人才培养模式上滞后于市场对人才的需求，仍可能以错配的形式造成高等教育溢价的降低。对于这一情形，应根据市场动态需求，及时改进大学生培养模式和培养内容，并正视职业教育不可或缺的互补作用。

户籍制度对不同学历劳动者的差异化限制可能对高等教育溢价和收入差距产生影响，相对于户籍放开的情形，对高学历劳动者偏好的户籍制度提升了大城市高学历劳动者相对比重，降低了大城市高等教育溢价。对于高学历劳动者而言，大城市更高的住房租金等支出对冲了可能存在的高等教育溢出效应，在不同规模城市生活的收支比差异不大。

1999 年及随后的大学扩招政策和 2009 年以后延续的扩招政策对就业产业的影响及其他方面的效应，仍需进一步研究。近期的研究生扩招能否缓解高校毕业生就业压力，也值得讨论。如何回归教育本质，避免“教育竞赛”，是中国的教育改革需要深入探讨的问题，也是本文后续的研究方向。

参考文献：

- 蔡昉(2018)，《农业劳动力转移潜力耗尽了么?》，《中国农村经济》第 9 期，第 2 - 13 页。
- 陈飞、苏章杰(2021)，《城市规模的工资溢价：来源与经济机制》，《管理世界》第 1 期，第 19 - 32 + 2 页。
- 陈技伟、冯帅章(2022)，《高校扩招如何影响中等职业教育?》，《经济学(季刊)》第 1 期，第 21 - 44 页。
- 方森辉、毛其淋(2021)，《高校扩招、人力资本与企业出口质量》，《中国工业经济》第 11 期，第 97 - 115 页。
- 胡雯、张锦华(2021)，《密度、距离与农民工工资：溢价还是折价?》，《经济研究》第 3 期，第 167 - 185 页。
- 梁文泉、陆铭(2016)，《后工业化时代的城市：城市规模影响服务业人力资本外部性的微观证据》，《经济研究》第 12 期，第 90 - 103 页。
- 刘生龙、胡鞍钢(2018)，《大学教育回报：基于大学扩招的自然实验》，《劳动经济研究》第 4 期，第 48 - 70 页。
- 刘生龙、周绍杰、胡鞍钢(2016)，《义务教育法与中国城镇教育回报率：基于断点回

- 归设计》，《经济研究》第2期，第154-167页。
- 陆铭、高虹、佐藤宏（2012），《城市规模与包容性就业》，《中国社会科学》第10期，第47-66页。
- 马光荣、纪洋、徐建炜（2017），《大学扩招如何影响高等教育溢价？》，《管理世界》第8期，第52-63页。
- 吴要武（2010），《寻找阿基米德的“杠杆”——“出生季度”是个弱工具变量吗？》，《经济学（季刊）》第2期，第661-686页。
- 吴要武、赵泉（2010），《高校扩招与大学毕业生就业》，《经济研究》第9期，第93-108页。
- 邢春冰、李实（2011），《扩招“大跃进”、教育机会与大学毕业生就业》，《经济学（季刊）》第4期，第1187-1208页。
- 张春泥、谢宇（2017），《入学年龄限制真的造成了“七八月陷阱”吗？——兼评刘德寰、李雪莲〈“七八月”的孩子们〉》，《社会学研究》第1期，第54-77页。
- 钟粤俊、陆铭、奚锡灿（2020），《集聚与服务业发展——基于人口空间分布的视角》，《管理世界》第11期，第35-49页。
- 周茂、李雨浓、姚星、陆毅（2019），《人力资本扩张与中国城市制造业出口升级：来自高校扩招的证据》，《管理世界》第5期，第64-77页。
- Acemoglu, Daron & David Autor (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In Orley Ashenfelter & David Card (eds.), *Handbook of Labor Economics (Volume 4)*. Amsterdam: Elsevier, pp. 1043-1171.
- Acemoglu, Daron & Joshua Angrist (2000). How Large are Human-capital Externalities? Evidence from Compulsory Schooling Laws. *NBER Macroeconomics Annual*, 15, 9-59.
- Acemoglu, Daron (1998). Why do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *Quarterly Journal of Economics*, 113 (4), 1055-1089.
- Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva & Valerio Sterzi (2018). Dancing with the Stars: Innovation Through Interactions. *NBER Working Paper*, No. 24466.
- Angrist, Joshua (2014). The Perils of Peer Effects, *Labour Economics*, 30, 98-108.
- Bianchi, Nicola (2020). The Indirect Effects of Educational Expansions: Evidence from a Large Enrollment Increase in University Majors. *Journal of Labor Economics*, 38 (3), 767-804.

- Borjas, George (2003). The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market. *Quarterly Journal of Economics*, 118 (4), 1335 – 1374.
- Camacho, Adriana, Julián Messina & Juan Barrera (2017). The Expansion of Higher Education in Colombia: Bad Students or Bad Programs? <https://ssrn.com/abstract=2921965>.
- Card, David & John DiNardo (2002). Skill-biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles. *Journal of Labor Economics*, 20 (4), 733 – 783.
- Card, David & Thomas Lemieux (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-based Analysis. *Quarterly Journal of Economics*, 116 (2), 705 – 746.
- Card, David (1999). The Causal Effect of Education on Earnings. In Orley Ashenfelter & David Card (eds.), *Handbook of Labor Economics (Volume 3)*. Amsterdam: Elsevier, pp. 1801 – 1863.
- Card, David (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, 19 (1), 22 – 64.
- Card, David (2009). Immigration and Inequality. *The American Economic Review: Papers & Proceedings*, 99 (2), 1 – 21.
- Card, David, Jörg Heining & Patrick Kline (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics*, 128 (3), 967 – 1015.
- Carneiro, Pedro & Sokbae Lee (2011). Trends in Quality-adjusted Skill Premia in the United States, 1960 – 2000. *The American Economic Review*, 101 (6), 2309 – 2349.
- Che, Yi & Lei Zhang (2018). Human Capital, Technology Adoption and Firm Performance: Impacts of China's Higher Education Expansion in the Late 1990s. *Economic Journal*, 128 (614), 2282 – 2320.
- Chen, Yi, Ziyang Fan, Xiaomin Gu & Lian Zhou (2020). Arrival of Young Talent: The Send-Down Movement and Rural Education in China. *The American Economic Review*, 110 (11), 3393 – 3430.
- Ciccone, Antonio & Giovanni Peri (2006). Identifying Human-Capital Externalities: Theory with Applications. *Review of Economic Studies*, 73 (2), 381 – 412.
- Duflo, Esther (2004). The Medium Run Effects of Educational Expansion: Evidence from a

- Large School Construction Program in Indonesia. *Journal of Development Economics*, 74 (1), 163 – 197.
- Feng, Shuaizhang & Xiaoyu Xia (2022). Heterogeneous Firm Responses to Increases in High-skilled Workers: Evidence from China's College Enrollment Expansion. *China Economic Review*, 73, <https://doi.org/10.1016/j.chieco.2022.101791>.
- Goldin, Claudia & Lawrence Katz (2010). *The Race Between Education and Technology*. Cambridge: Harvard University Press.
- Katz, Lawrence & Kevin Murphy (1992). Changes in Relative Wages, 1963 – 1987: Supply and Demand Factors. *Quarterly Journal of Economics*, 107 (1), 35 – 78.
- Li, Hongbin, Yueyuan Ma, Lingsheng Meng, Xue Qiao & Xinzheng Shi (2017). Skill Complementarities and Returns to Higher Education: Evidence from College Enrollment Expansion in China. *China Economic Review*, 46, 10 – 26.
- Ma, Xiao (2020). College Expansion, Trade and Innovation: Evidence from China. *MPRA Paper*, No. 109469.
- Manacorda, Marco, Alan Manning & Jonathan Wadsworth (2012). The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain. *Journal of the European Economic Association*, 10 (1), 120 – 151.
- Meng, Xin (2012). Labor Market Outcomes and Reforms in China. *Journal of Economic Perspectives*, 26 (4), 75 – 102.
- Meng, Xin, Kailing Shen & Sen Xue (2013). Economic Reform, Education Expansion, and Earnings Inequality for Urban Males in China, 1988 – 2009. *Journal of Comparative Economics*, 41 (1), 227 – 244.
- Moretti, Enrico (2004). Workers' Education, Spillovers, and Productivity: Evidence from Plant-level Production Functions. *The American Economic Review*, 94 (3), 656 – 690.
- Ottaviano, Gianmarco & Giovanni Peri (2012). Rethinking the Effect of Immigration on Wages. *Journal of the European Economic Association*, 10 (1), 152 – 197.
- Romer, Paul (1990). Endogenous Technological Change. *Journal of Political Economy*, 98 (5, Part 2), S71 – S102.
- Song, Jae, David Price, Fatih Guvenen, Nicholas Bloom & Till Wachter (2019). Firming up Inequality. *Quarterly Journal of Economics*, 134 (1), 1 – 50.
- Vytlacil, Edward (2002). Independence, Monotonicity, and Latent Index Models: An

Equivalence Result. *Econometrica*, 70 (1), 331 – 341.

Zhang, Junsen (2017). The Evolution of China's One-child Policy and Its Effects on Family Outcomes. *Journal of Economic Perspectives*, 31 (1), 141 – 160.

Zhang, Junsen, Yaohui Zhao, Albert Park & Xiaoqing Song (2005). Economic Returns to Schooling in Urban China, 1988 to 2001. *Journal of Comparative Economics*, 33 (4), 730 – 752.

New Trends and Explanations of Higher Education Premium: An Analysis for Chinese Migrants

Yan Weibo

(School of Public Finance and Taxation, Zhongnan University of Economics and Law)

Abstract: This study finds a significant decline in the college premium by examining data from the China Migrants Dynamic Survey from 2011 to 2017. The wage premium of college-educated laborers, compared to those with middle and high school education, declines from 8.18 percent to 4.65 percent during the period. An instrumental variable is constructed to deal with endogeneity of the relative supply of college-educated laborers. Analysis on the substitution elasticity between college-educated laborers and high school equivalents shows that a 10 percent increase in college-educated labor supply is associated with a decline of the college premium by 1.3 percent to 3 percent. The surplus of the college-educated laborers at least accounts for more than half of the decline of the college premium. A possible spillover effect of higher education is also discussed. Even if college-educated laborers enjoy higher college premiums in cities with higher proportion of well-educated, they have to face higher housing costs. As a result, the income-expenditure ratio is similar across cities of different scales. This paper is expected to invoke more discussions in higher education expansion, education investment, and migration decision.

Keywords: higher education supply, return to education, migrants, spillover effect

JEL Classification: I23, I24, I26

(责任编辑:一帆)