**大学教育的社会回报**

——扩招与人力资本外溢效应

丁相元 张子尧 黄炜

**目录**

[附录I 制度背景 1](#_Toc161668789)

[附录II 样本清理和描述性统计 2](#_Toc161668790)

[附录III 人力资本外溢效应存在的定性证据 3](#_Toc161668791)

[附录IV 稳健性检验 5](#_Toc161668792)

[附录V 人力资本外溢效应的潜在机制：非参数动态双重差分估计结果 1](#_Toc161668793)1

[附录VI 关于大学教育回报的进一步讨论：分析框架和估计结果 1](#_Toc161668794)3

附录I 制度背景

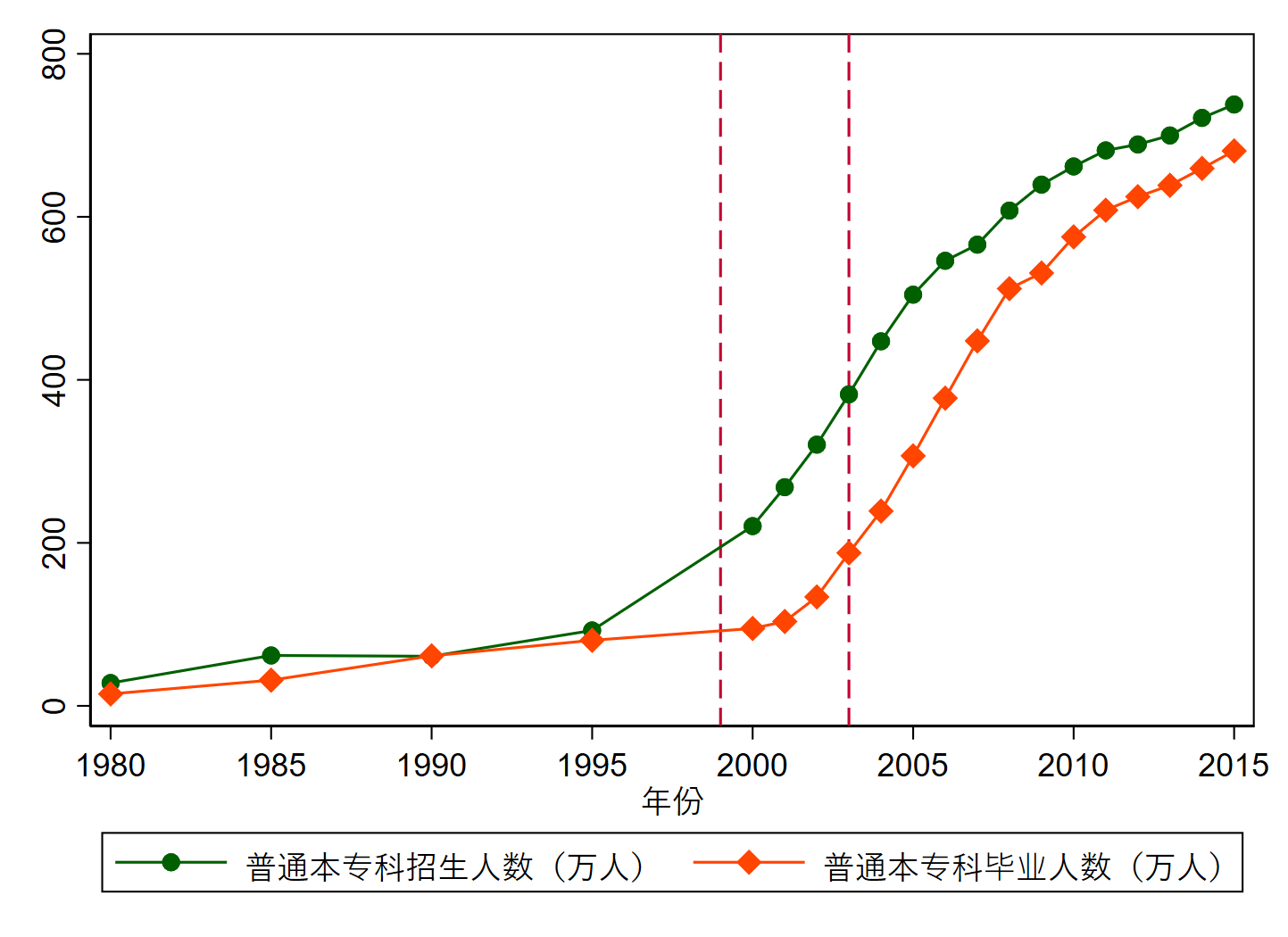


图 I1 1980-2015年中国高校历年招生和毕业人数趋势

注：数据来源为《中国统计年鉴》。左侧虚线为高校扩招政策实施的时间（1999年），右侧虚线为高校扩招政策实施后，大学毕业生陆续进入劳动力市场的时间（2003年）。

附录II 样本清理和描述性统计

本文按照如下步骤对UHS原始数据进行清理：（1）剔除20周岁以下和60周岁以上样本；（2）运用年度CPI对收入、消费等名义变量进行平减；（3）在0.1%水平上对收入、消费等连续型变量进行缩尾处理。经过清理后，最终得到275,998个观测值。

表II1汇报了本文所使用的主要数据和变量的描述性统计，其中面板A为UHS 1999—2007年的样本，面板B为工商注册企业信息2000—2007年的样本，面板C为城市统计年鉴1999—2007年的样本。

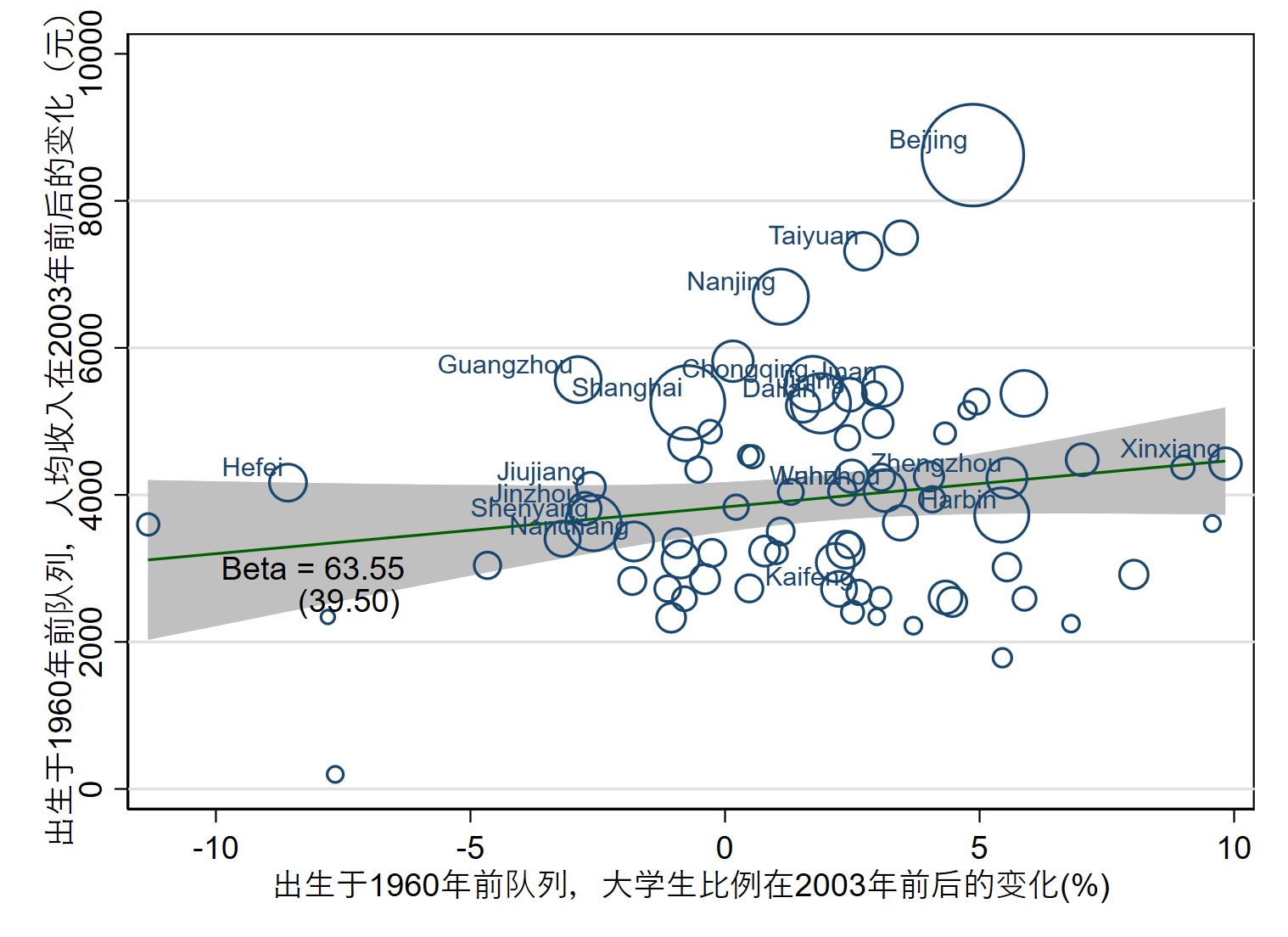
表II1 样本描述性统计

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 变量 | 观测值 | 平均值 | 标准差 | 最小值 | 最大值 |
| **面板A: UHS样本（个人-年层面）** | |  |  |  |  |
| 出生年份 | 275998 | 1962.45 | 10.47 | 1940 | 1987 |
| 大学毕业生 (是=1)[[1]](#footnote-0) | 275998 | 0.29 | 0.455 | 0 | 1 |
| 男性 (是=1) | 275998 | 0.49 | 0.500 | 0 | 1 |
| 个人收入 (元)[[2]](#footnote-1) | 275998 | 12628.98 | 11558.36 | 0 | 63160.00 |
| 家庭人均收入 (元) | 275735 | 10584.85 | 7785.81 | 138.39 | 44113.89 |
| 从事第三产业 (是=1) | 275998 | 0.37 | 0.481 | 0 | 1 |
| **面板B:工商注册企业数据（城市-年层面）** | | | | | |
| 各产业新增加雇员数 |  |  |  |  |  |
| 所有产业 | 1496 | 120.77 | 196.414 | 1.70 | 2164.35 |
| 第一产业 | 1496 | 2.23 | 11.932 | 0.00 | 366.80 |
| 第二产业 | 1496 | 40.73 | 67.958 | 0.53 | 641.20 |
| 第三产业 | 1496 | 120.77 | 196.414 | 1.70 | 2164.35 |
| **面板C: 城市统计年鉴数据（城市-年层面）** | | | |  |  |
| 人均固定资产投资 (元) | 1551 | 5790.32 | 7560.62 | 171.83 | 66163.02 |
| 人均工业产值 (元) | 1549 | 20986.40 | 42318.96 | 194.85 | 651310.94 |
| 各产业从业人员比重 (%) | |  |  |  |  |
| 第一产业 | 1549 | 4.74 | 8.298 | 0.03 | 63.60 |
| 第二产业 | 1550 | 43.88 | 12.740 | 11.64 | 81.58 |
| 第三产业 | 1550 | 51.39 | 12.559 | 17.89 | 86.94 |
| 各产业从业人员人数 (万人) | |  |  |  |  |
| 第一产业 | 1549 | 17.24 | 30.257 | 0.04 | 441.95 |
| 第二产业 | 1550 | 187.47 | 144.430 | 9.09 | 1519.63 |
| 第三产业 | 1550 | 239.27 | 188.197 | 2.87 | 1690.13 |

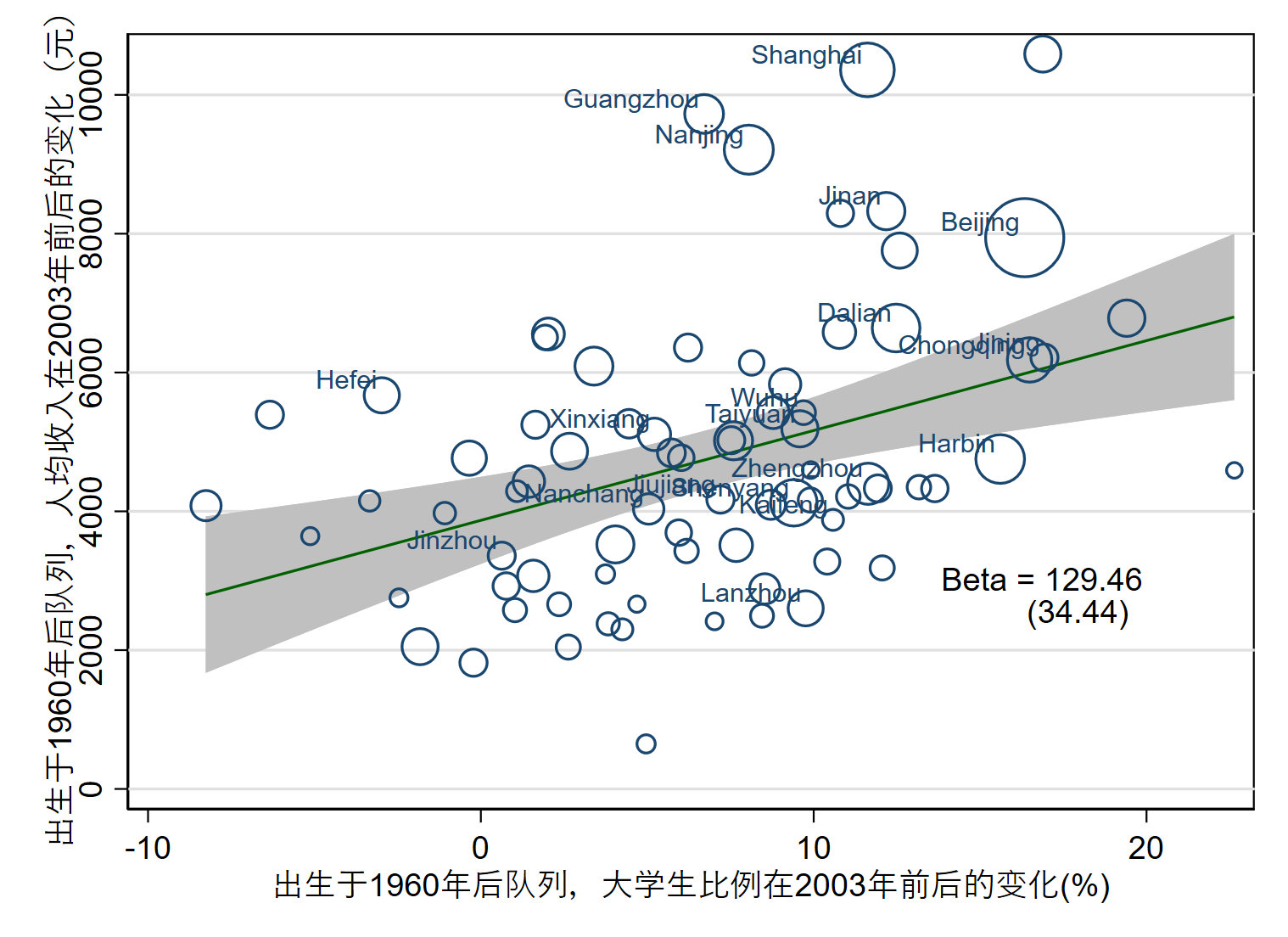
注：面板A样本来自UHS 1999—2007年数据，面板B样本来自2000—2007年工商注册企业数据，面板C样本来自1999—2007年城市统计年鉴数据。

附录III 人力资本外溢效应存在的定性证据

正如正文提到的，本文主要通过分析扩招政策在不同城市中影响的差异性，来分析和找出人力资本外溢效应存在的潜在证据。因此，我们在每个城市样本中，首先计算了每个出生队列在2003年前后大学毕业生比例和人均收入的变化情况，然后就两者的相关关系进行汇报（每个坐标轴均反映了在城市-队列层面，该变量在2003至2007年与1999至2002年的变化值）。图III1中的a图展示的是，每个城市出生于1960年前群体中，2003年前后的人均收入和大学毕业生比例变化的相关关系。可以发现，其一，该群体中大学毕业生比例变化并不明显（均值在0左右，范围主要在-5%~+5%之间），与正文发现一致；其二，该群体中人均收入和大学生毕业生比例之间的相关关系并不显著（估计系数约为64，且在10%的水平下不显著）。上述结果说明，出生于1960年前群体的收入增长并不能由他们自身人力资本的变化来解释。图III1中的b图则汇报了出生于1960年后群体的结果。可以发现，其一，和正文发现一致的是，大学毕业生比例变化明显（均值在10%左右，范围主要在-5%~+20%之间）；其二，该群体中人均收入和大学生毕业生比例之间具有显著的相关关系（估计系数约为129，且在1%的水平下显著）。该结果说明，出生于1960年后群体的收入增长，可能部分可以被他们自身人力资本的提升来解释。



a. 出生于1960前队列，大学生比例和人均收入相关关系



b. 出生于1960后队列，大学生比例和人均收入相关关系

图 III1 队列内部的大学比例和人均收入相关关系

注：数据来源为UHS 1999-2007年样本。图中圆圈与样本数据中的城市相对应，圆圈的面积大小与城市的样本量大小为正相关关系。由于图中城市数较多，我们仅选择了部分直辖市和省会城市等高校数量较多、受扩招政策影响效果更为明显的大城市进行标注。

进一步，我们尝试对出生于1960年前群体的收入增长进行解释，并验证人力资本在跨出生队列的外溢效应。图III2报告了每个城市出生于1960年前群体的人均收入与和该城市出生于1960年后群体的大学毕业生比例变化之间的相关性。结果显示，二者呈现显著的正相关关系（估计系数约为82，且在1%的水平下显著）。基于此，我们可以初步推断扩招政策带来的人力资本提高会在出生队列之间产生外溢效应，即年轻群体的教育增长能够促进年长群体的收入提高。但是，上述讨论仅提供了相关性的定性证据，并未考虑到不同城市随时间变化的特征影响以及其他可能混淆因素。因此，在正文分析中，我们利用高校扩招政策提供的准自然实验框架，建立经典双重差分及动态双重差分模型对人力资本外溢效应的存在现象和潜在机制进行分析和讨论。

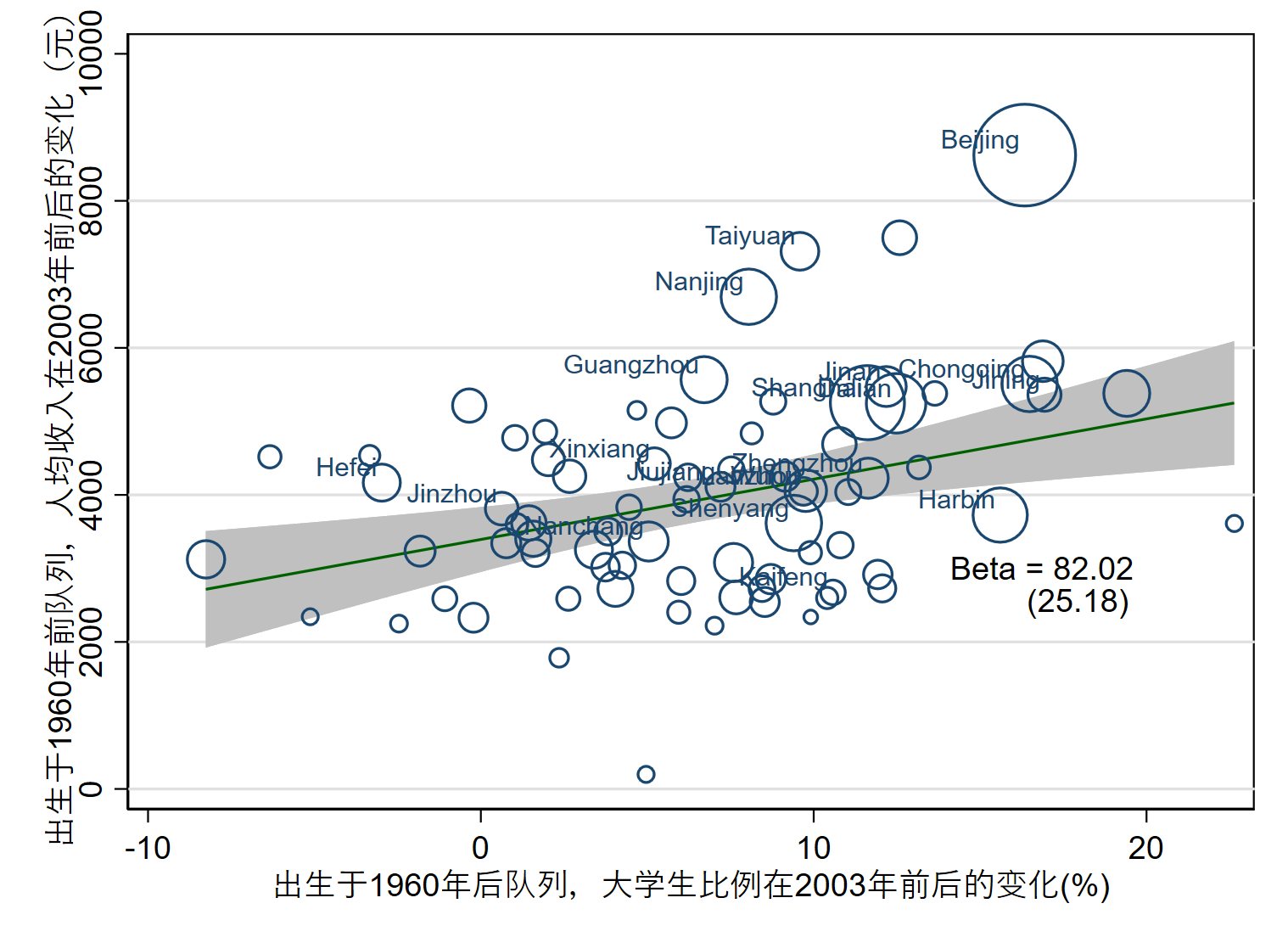


图 III2 出生于1960年后队列的大学比例和出生于1960年前队列的人均收入相关关系

注：数据来源为UHS 1999-2007年样本。图中圆圈与样本数据中的城市相对应，圆圈的面积大小与城市的样本量大小为正相关关系。由于图中城市数较多，我们仅选择了部分直辖市和省会城市等高校数量较多、受扩招政策影响效果更为明显的大城市进行标注。

附录IV 稳健性检验

**（1）排除家庭内部溢出效应的渠道**

对于正文中发现出生于1960年前群体的个人收入在2003年后出现增加的情况，除跨出生队列间存在人力资本外溢效应的解释外，还有一个可能的解释是家庭内部存在人力资本溢出效应，即子代的人力资本提高不仅能通过人力资本积累效应促进子代自身的个人收入提高，同时也可以通过对父代的赡养转移支付或知识信息传递等方式促进父代的个人收入提高[[3]](#footnote-2)。

为了检验上述关于家庭内部溢出效应的解释是否成立，我们将总样本分为家中最年轻的早于1960年出生和家中最年轻的晚于1960年出生两个样本分别进行估计，就前者而言，家户内所有成员在扩招政策实施时均已完成了教育，因此既不存在由于人力资本积累效应导致个人收入改善的渠道，也不存在家庭内部溢出效应导致个人收入增加的可能。

表IV1汇报了采用经典双重差分和参数动态双重差分模型进行定量估计的结果，可以发现，即使对于那些人力资本积累效应和家庭内部溢出效应均不存在的家庭而言（即家中最年轻的早于1960年出生的样本），他们的家户人均收入也会在2003年后出现显著的相对提高（约1,200元）。因此我们可以排除家庭内部溢出效应这一可能的解释，说明扩招政策主要是通过在跨出生队列间带来人力资本外溢效应，进而使得出生于1960年前群体的个人收入获得提高。此外，我们还可以发现，基于家户人均收入的估计结果与正文对个人收入的估计结果基本一致，证明了本文估计效果的稳健性。

表 IV1 扩招政策对家户人均收入的影响

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| 因变量 | 家户人均收入（元） | | | | | |
| 样本 | 全样本 | 最年长>= 1960 | 最年轻 <= 1960 | 全样本 | 最年长>= 1960 | 最年轻 <= 1960 |
| **面板A: 经典双重差分估计** | | | | | | |
| (大学数量指标） | 44.42 | 2.411 | 122.6 | 2.313 | -0.0648 | 4.599 |
| \* (年份=2002) | (218.3) | (211.0) | (252.1) | (5.913) | (6.165) | (7.232) |
| (大学数量指标） | 1,316\*\*\* | 1,245\*\*\* | 1,226\*\*\* | 57.15\*\*\* | 53.98\*\*\* | 53.71\*\*\* |
| \* (年份>=2003) | (384.0) | (371.0) | (419.7) | (4.376) | (4.791) | (6.079) |
|  |  |  |  |  |  |  |
| **面板B: 动态双重差分估计，大学数量指标\*年份，2001年以及以前为参照组** | | | | | | |
| 2002 | 40.83 | 0.233 | 118.1 | 2.496 | 0.157 | 4.749 |
|  | (218.7) | (211.2) | (252.4) | (5.887) | (6.153) | (7.211) |
| 2003 | 570.6\*\* | 533.0\*\* | 566.4\* | 20.99\*\*\* | 18.90\*\*\* | 19.73\*\* |
|  | (244.8) | (237.3) | (288.5) | (5.716) | (6.324) | (9.019) |
| 2004 | 797.7\*\*\* | 739.3\*\*\* | 753.3\*\*\* | 36.23\*\*\* | 32.69\*\*\* | 33.36\*\*\* |
|  | (269.0) | (257.4) | (281.7) | (4.408) | (4.154) | (4.670) |
| 2005 | 1,080\*\*\* | 972.5\*\* | 1,006\*\* | 51.57\*\*\* | 48.47\*\*\* | 49.88\*\*\* |
|  | (399.9) | (390.8) | (448.1) | (5.244) | (4.814) | (5.153) |
| 2006 | 1,703\*\*\* | 1,624\*\*\* | 1,667\*\*\* | 72.21\*\*\* | 68.17\*\*\* | 69.53\*\*\* |
|  | (540.4) | (525.8) | (602.9) | (5.783) | (6.293) | (6.851) |
| 2007 | 2,322\*\*\* | 2,220\*\*\* | 2,334\*\*\* | 85.90\*\*\* | 82.03\*\*\* | 85.85\*\*\* |
|  | (606.6) | (599.0) | (737.1) | (5.746) | (6.100) | (7.696) |
|  |  |  |  |  |  |  |
| **面板C: 动态双重差分估计平均效果** | | | | | | |
| 2003-2007 | 1,295\*\*\* | 1,218\*\*\* | 1,265\*\*\* | 53.38\*\*\* | 50.05\*\*\* | 51.67\*\*\* |
|  | (374.6) | (359.9) | (427.4) | (4.220) | (4.619) | (6.063) |
| 城市和年份固定效应 | 是 | 是 | 是 | 是 | 是 | 是 |
| 性别、年龄及交互项 | 是 | 是 | 是 | 是 | 是 | 是 |
| 样本量 | 275,735 | 224,326 | 169,158 | 275,735 | 224,326 | 169,158 |

注：数据来源为UHS 1999-2007年样本。面板A结果为根据正文模型（1）进行估计得到，面板B结果为根据正文模型（3）进行估计得到，面板C结果为在正文模型（3）基础上、计算平均效果得到。其中第（1）-（3）列中的处理变量为根据城市1998年高等院校数构建得到的二元分组变量，第（4）-（6）列中的处理变量为根据城市1998年高等院校数构建得到的连续强度变量。括号内是聚类在城市层面的稳健性标准误，\*、\*\*、\*\*\*分别表示在10%、5%、1%的水平上显著。

**（2）事前平行趋势检验**

在本文中，我们主要使用双重差分法来对扩招政策在不同出生队列中的影响进行分析，其关键假定在于假定扩招政策冲击没有发生时，处理组和控制组具有良好的平行趋势。为了对该假定进行更加充分的验证，我们将分析样本期扩展为1996—2007年，并基于该样本进行事前平行趋势检验。从汇报在附录附图三中的非参数动态双重差分估计结果可以发现，在2001年之前，无论是从每个年份系数估计值的显著性还是各个系数估计值的联合显著性看，处理组城市和控制组城市之间的事前平行趋势均满足，进一步支持了平行趋势假设这一核心假定的成立。

此外，我们也将基于1996—2007年样本估计得到的动态双重差分估计平均效果分别汇报在表IV2和表IV3的面板A中。结果显示，扩招政策对大学毕业生比例和个人收入的影响仍与正文结论保持一致。

**（3）控制个体特征和城市特征变量**

为尽可能缓解遗漏重要变量引起的估计偏差，一方面，考虑到个体户籍和工作行业均是影响个体工资收入的重要因素（张原和陈建奇，2008；宁光杰和段乐乐，2017），我们在上文模型中相应控制了个体户籍和行业特征的固定效应，就扩招政策对大学毕业生比例和个人收入的影响再次进行估计；另一方面，由于部分城市层面的宏观经济变量在处理组城市和控制组城市之间的发展趋势差异，可能也会对两组城市间个体收入的相对增长产生影响，我们相应地在模型中引入了1999年城市特征与年份固定效应的交互项，并再次进行估计。结合已有文献（周茂等，2019；葛润和黄家林，2020），我们控制的城市特征包括经济总量（用GDP表示）、人口规模（用常住人口数表示）、产业结构（用第一、二产业占比表示）、劳动力市场状况（用城镇登记失业人口表示）、金融发展水平（用金融从业人数占比表示）和基础设施建设（用城市公共图书馆藏书量和医院总数表示）等方面。我们将上述估计得到的动态双重差分估计平均效果分别汇报在表IV2和表IV3的面板B、面板C中。

通过比较可以发现，在分别控制个体特征和城市特征等潜在遗漏变量的影响后，无论是从经济显著性还是从统计显著性看，扩招政策对大学毕业生比例和个人收入影响的估计结果均与上文结论基本相符，证实了本文关于人力资本外溢效应存在的结论是稳健的。

**（4）控制地区特征的时间趋势**

在本文使用双重差分模型进行估计的过程中，如果不同地区教育或收入增长的时间趋势存在差异，那么这种趋势项的差异可能会使估计结果产生偏误。对于该问题，我们尝试在模型中引入省份和年份的交互项，以尽可能控制不同地区时间趋势的差异。根据汇报在面板D中的估计结果可以发现，在控制省份和年份的交互项后，估计得到的结果与基准结论基本一致。说明不同地区特征变化的时间趋势并不会对本文结论产生过多影响，证实了本文结论的稳健性和可靠性。

**（5）排除同期其他政策的影响**

本文主要是依赖于高校扩招政策这一外生冲击提供的准自然实验进行分析框架的建立和实证结果的分析，但在同时期内还可能存在其他政策会对劳动力市场和宏观经济产生冲击，进而使得本文的估计结果出现偏误。在2003年前后，最可能对本文结果产生影响的应为中国于2001年末加入WTO这一事件，影响渠道包括产品关税下降效应、进出口规模经济效应等（周茂等，2019）。

对此，我们在上文模型中引入城市层面关税变量和进出口变量作为控制变量后进行估计，相关结果分别报告在面板E中。通过比较可以发现，在引入关税变量和进出口变量作为控制变量后，估计结果与上文结论保持一致，因此可以基本排除中国加入WTO这一事件对本文估计结果的影响。

**（6）更换处理组城市的划分指标**

在本文使用的经典双重差分和动态双重差分模型中，处理组城市划分指标的不同会对估计结果产生直接影响。因此，我们也尝试直接使用各城市在1998年每万人大学生比例的中位数对城市进行分组，相关结果报告在面板F中。通过比较可以发现，在更换处理组城市的划分指标后，扩招政策对大学毕业生比例和个人收入影响的整体方向仍与上文结论保持一致，这也从处理组城市划分指标的角度证实了本文结论的可靠性。

**（7）更换出生队列的划分标准**

本文主要通过分析扩招政策对不同出生队列（特别是政策实施时就已完成教育的群体）的影响，进而证实了人力资本外溢效应的存在。因此，如何合理划分出生队列对本文研究结论有十分重要的影响。在正文中，我们就选择1960年作为划分标准的原因进行了阐释，并验证了该划分方式的有效性。在此基础上，我们分别将划分标准从1960年更换为1955年和1965年后再次估计，相关结果分别报告在面板G和面板H中。估计结果显示，在更换指标年份后，估计得到的结果均与正文结论一致，说明本文结论并不依赖于1960年这一划分标准的选择，进一步支持了本文结论的稳健性。

**（8）其他可能问题**

其一，考虑到在UHS的不同调查年份，本文所关注的出生于1960年前群体和出生于1960年后群体的年龄组成也在发生改变，这些差异也可能会对我们的估计产生影响。为了保持样本年龄组成在不同调查年份的一致性，对于出生于1960年前队列，我们仅保留了每年47-60岁的群体；对于出生于1960年后队列，我们则保留了每年20-39岁的样本。此外，考虑到大学生毕业年龄通常为21-22岁以及女性退休年龄一般为50岁或55岁，我们进一步将样本限制为22岁及以上、女性50岁及以下群体，以尽可能保证样本处于劳动力市场中。基于该限制年龄后样本的估计结果汇报在面板I中。

其二，在扩招政策产生影响的过程中，由于人力资本的集聚效应以及劳动力市场一般均衡效应的存在，可能会出现不同地区之间劳动力发生流动的情况。为了验证劳动力流动对估计结果的影响程度，我们将样本限制在没有进行跨城市流动的本地样本中并再次进行估计，相应结果汇报在面板J中。

通过比较可以发现，在对其他可能问题进行处理和分析后，本文结论依然保持稳健。

表 IV2 稳健性检验：扩招政策对大学毕业生比例的影响

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| 变量 | 大学毕业生比例 （%） | | | | | | |
| 样本 | 全样本 | 生于指标年份后 | 生于指标年份前 | 全样本 | 生于指标年份后 | 生于指标年份前 |
| **面板A: 换用UHS 1996-2007年样本** | | | | | | | |
| 2002 | -0.238 | 0.890 | -0.408 | -0.0189 | -0.00414 | -0.00677 |
|  | (1.037) | (1.443) | (0.838) | (0.0230) | (0.0316) | (0.0191) |
| 2003-2007 | 1.994\*\* | 5.080\*\*\* | -0.744 | 0.0541\*\* | 0.142\*\*\* | -0.00289 |
|  | (0.999) | (1.249) | (0.915) | (0.0232) | (0.0226) | (0.0327) |
|  |  |  |  |  |  |  |
| **面板B: 控制个体户籍和工作行业的固定效应** | | | | | | | |
| 2002 | -0.350 | 1.235 | -0.952 | -0.00757 | 0.0239 | -0.00248 |
|  | (0.845) | (1.106) | (0.818) | (0.0161) | (0.0236) | (0.0179) |
| 2003-2007 | 1.124 | 3.795\*\*\* | -1.293 | 0.0482\*\*\* | 0.118\*\*\* | 0.00410 |
|  | (0.974) | (1.309) | (0.944) | (0.0157) | (0.0249) | (0.0217) |
|  |  |  |  |  |  |  |
| **面板C: 控制1999年城市特征交互年份固定效应** | | | | | | | |
| 2002 | 0.303 | 2.012 | -0.688 | -0.0439 | 0.0161 | -0.0604 |
|  | (1.247) | (1.590) | (1.079) | (0.0617) | (0.0817) | (0.0542) |
| 2003-2007 | 1.521 | 4.419\*\* | -1.111 | 0.0404 | 0.192\* | -0.0813 |
|  | (1.363) | (1.694) | (1.216) | (0.0807) | (0.105) | (0.0620) |
|  |  |  |  |  |  |  |
| **面板D: 控制省份和年份的交互项** | | | | | | | |
| 2002 | 0.117 | 1.237 | -0.316 | -0.00988 | -0.0203 | 0.00946 |
|  | (1.085) | (1.355) | (1.090) | (0.0544) | (0.0679) | (0.0606) |
| 2003-2007 | 1.993\* | 4.643\*\*\* | -0.572 | 0.107\* | 0.200\*\* | -0.00961 |
|  | (1.135) | (1.454) | (1.015) | (0.0630) | (0.0836) | (0.0482) |
|  |  |  |  |  |  |  |
| **面板E: 排除同期WTO政策影响** | | | | | | | |
| 2002 | 0.0756 | 0.978 | 0.0305 | -0.00294 | -0.0168 | 0.0299 |
|  | (1.053) | (1.402) | (0.973) | (0.0219) | (0.0281) | (0.0208) |
| 2003-2007 | 1.949\* | 4.729\*\*\* | -0.370 | 0.0618\*\* | 0.119\*\*\* | 0.0311 |
|  | (1.075) | (1.342) | (1.061) | (0.0261) | (0.0272) | (0.0359) |
|  |  |  |  |  |  |  |
| **面板F: 换用每万人大学生比例划分处理组城市** | | | | | | | |
| 2002 | -2.025 | -1.354 | -1.400 | -0.0101 | -0.0131 | -0.00214 |
|  | (1.306) | (1.742) | (1.166) | (0.00684) | (0.00957) | (0.00760) |
| 2003-2007 | -0.0490 | 2.955\* | -2.096 | 0.00426 | 0.0194 | -0.00500 |
|  | (1.285) | (1.689) | (1.263) | (0.00969) | (0.0120) | (0.00996) |
|  |  |  |  |  |  |  |
| **面板G: 换用1955年划分出生队列** | | |  |  |  |  |
| 2002 | 0.118 | 0.895 | -0.415 | -0.00140 | 0.0177 | -0.00316 |
|  | (1.001) | (1.188) | (0.896) | (0.0210) | (0.0257) | (0.0150) |
| 2003-2007 | 2.281\*\* | 4.135\*\*\* | 0.0278 | 0.0687\*\*\* | 0.134\*\*\* | 0.00657 |
|  | (1.126) | (1.310) | (1.126) | (0.0253) | (0.0231) | (0.0238) |
|  |  |  |  |  |  |  |
| **面板H: 换用1965年划分出生队列** | | | | | | | |
| 2002 | 0.118 | -1.265 | 0.685 | -0.00140 | -0.105\*\*\* | 0.0316 |
|  | (1.001) | (1.678) | (1.015) | (0.0210) | (0.0316) | (0.0274) |
| 2003-2007 | 2.281\*\* | 3.072\*\* | 0.919 | 0.0687\*\*\* | 0.0614\* | 0.0355 |
|  | (1.126) | (1.507) | (1.173) | (0.0253) | (0.0337) | (0.0395) |
|  |  |  |  |  |  |  |
| **面板I: 保留限制年龄后样本** | | | | | | | |
| 2002 | 0.0593 | 1.359 | -0.327 | -0.0320 | 0.0109 | -0.0272 |
|  | (1.239) | (1.502) | (1.371) | (0.0224) | (0.0291) | (0.0245) |
| 2003-2007 | 3.004\*\* | 7.518\*\*\* | -1.620 | 0.0623\*\* | 0.224\*\*\* | -0.0499 |
|  | (1.267) | (1.551) | (1.601) | (0.0287) | (0.0388) | (0.0466) |
|  |  |  |  |  |  |  |
| **面板J: 保留没有进行跨城市流动的本地样本** | | | | | | | |
| 2002 | 0.542 | 1.967 | 0.0667 | 0.0209 | 0.0433 | 0.0280 |
|  | (1.001) | (1.317) | (0.923) | (0.0221) | (0.0276) | (0.0200) |
| 2003-2007 | 2.739\*\* | 5.892\*\*\* | -0.226 | 0.0865\*\*\* | 0.165\*\*\* | 0.0336 |
|  | (1.157) | (1.434) | (1.113) | (0.0240) | (0.0289) | (0.0346) |
|  |  |  |  |  |  |  |
| 城市和年份固定效应 | 是 | 是 | 是 | 是 | 是 | 是 |
| 性别、年龄及交互项 | 是 | 是 | 是 | 是 | 是 | 是 |

注：面板A数据来源为UHS 1996-2007年样本，其他面板数据来源为UHS 1999-2007年样本。其中第（1）-（3）列中的处理变量为根据城市1998年高等院校数构建得到的二元分组变量，第（4）-（6）列中的处理变量为根据城市1998年高等院校数构建得到的连续强度变量。括号内是聚类在城市层面的稳健性标准误，\*、\*\*、\*\*\*分别表示在10%、5%、1%的水平上显著。

表 IV3 稳健性检验：扩招政策对个人收入的影响

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| 变量 | 个人收入 （元） | | | | | |
| 样本 | 全样本 | 生于指标年份后 | 生于指标年份前 | 全样本 | 生于指标年份后 | 生于指标年份前 |
| **面板A: 换用UHS 1996-2007年样本** | | | | | | |
| 2002 | 644.9\*\*\* | 389.7 | 770.1\*\*\* | 24.59\*\*\* | 18.33\*\* | 25.72\*\*\* |
|  | (241.2) | (271.8) | (246.3) | (4.964) | (8.108) | (4.119) |
| 2003-2007 | 1,954\*\*\* | 1,764\*\*\* | 1,981\*\*\* | 77.27\*\*\* | 73.60\*\*\* | 75.70\*\*\* |
|  | (548.3) | (535.4) | (641.5) | (6.861) | (15.80) | (11.56) |
|  |  |  |  |  |  |  |
| **面板B: 控制个体户籍和工作行业的固定效应** | | | | | | |
| 2002 | 160.5 | 154.9 | 124.6 | 4.201 | 5.464 | 1.241 |
|  | (252.2) | (294.4) | (237.4) | (6.301) | (10.17) | (3.545) |
| 2003-2007 | 1,333\*\*\* | 1,166\*\* | 1,356\*\*\* | 58.79\*\*\* | 54.17\*\*\* | 56.73\*\*\* |
|  | (495.6) | (498.9) | (503.8) | (9.016) | (14.74) | (7.305) |
|  |  |  |  |  |  |  |
| **面板C: 控制1999年城市特征交互年份固定效应** | | | | | | |
| 2002 | 277.4 | 142.6 | 379.1 | 18.95 | 13.54 | 21.31 |
|  | (254.1) | (292.5) | (255.2) | (20.18) | (25.99) | (17.08) |
| 2003-2007 | 1,161\*\*\* | 992.2\*\* | 1,370\*\*\* | 84.07\*\*\* | 73.37\* | 96.70\*\*\* |
|  | (432.1) | (484.2) | (405.4) | (29.86) | (38.13) | (23.38) |
|  |  |  |  |  |  |  |
| **面板D: 控制省份和年份的交互项** | | | | | | |
| 2002 | 245.3 | 41.53 | 395.7\* | 23.70\*\* | 9.224 | 31.60\*\* |
|  | (185.1) | (183.0) | (211.1) | (11.45) | (11.44) | (12.16) |
| 2003-2007 | 1,449\*\*\* | 1,400\*\*\* | 1,463\*\*\* | 99.28\*\*\* | 98.42\*\*\* | 99.51\*\*\* |
|  | (299.9) | (337.6) | (298.8) | (18.71) | (23.02) | (15.97) |
|  |  |  |  |  |  |  |
| **面板E: 排除同期WTO政策影响** | | | | | | |
| 2002 | 340.4 | 62.29 | 540.9\*\* | 11.17\*\* | 0.657 | 15.69\*\*\* |
|  | (256.7) | (299.7) | (251.2) | (5.401) | (7.837) | (4.137) |
| 2003-2007 | 1,655\*\*\* | 1,396\*\*\* | 1,737\*\*\* | 65.60\*\*\* | 56.19\*\*\* | 66.82\*\*\* |
|  | (472.1) | (456.9) | (540.3) | (6.999) | (14.65) | (11.11) |
|  |  |  |  |  |  |  |
| **面板F: 换用1952年高校学生数划分处理组城市** | | | | | | |
| 2002 | -39.11 | -238.4 | -30.65 | -67.82 | -184.5 | -16.01 |
|  | (240.3) | (291.7) | (214.1) | (183.2) | (217.7) | (163.6) |
| 2003-2007 | 1,826\*\*\* | 1,587\*\*\* | 1,722\*\* | 65.60\*\*\* | 58.65\*\*\* | 65.64\*\*\* |
|  | (543.9) | (533.8) | (650.3) | (7.348) | (14.56) | (11.23) |
|  |  |  |  |  |  |  |
| **面板G: 换用每万人大学生比例划分处理组城市** | | | | | | |
| 2002 | -123.8 | -303.5 | -56.38 | 1.888 | -0.00896 | 2.729\*\* |
|  | (237.4) | (271.4) | (253.3) | (1.546) | (2.182) | (1.325) |
| 2003-2007 | 1,107\*\* | 816.5\* | 1,248\*\* | 14.09\*\*\* | 11.10\*\*\* | 16.23\*\*\* |
|  | (461.5) | (457.9) | (521.7) | (3.602) | (3.547) | (4.647) |
|  |  |  |  |  |  |  |
| **面板H: 换用1955年划分出生队列** | | | | | | |
| 2002 | 234.7 | 183.9 | 303.5 | 7.815 | 7.746 | 4.741 |
|  | (236.0) | (257.3) | (241.8) | (4.833) | (5.691) | (4.121) |
| 2003-2007 | 1,545\*\*\* | 1,450\*\*\* | 1,689\*\*\* | 62.32\*\*\* | 61.72\*\*\* | 57.72\*\*\* |
|  | (478.7) | (483.9) | (501.9) | (6.618) | (9.680) | (9.400) |
|  |  |  |  |  |  |  |
| **面板I: 换用1965年划分出生队列** | | | | | | |
| 2002 | 234.7 | -178.3 | 315.3 | 7.815 | -12.29 | 10.96\*\*\* |
|  | (236.0) | (319.4) | (223.6) | (4.833) | (8.922) | (3.708) |
| 2003-2007 | 1,545\*\*\* | 1,124\*\* | 1,555\*\*\* | 62.32\*\*\* | 44.95\*\*\* | 63.14\*\*\* |
|  | (478.7) | (443.3) | (535.3) | (6.618) | (16.49) | (8.770) |
|  |  |  |  |  |  |  |
| **面板J: 保留限制年龄后样本** | | | | | | |
| 2002 | 209.6 | 192.8 | 239.9 | 3.992 | 6.254 | -0.211 |
|  | (302.3) | (356.1) | (287.7) | (8.414) | (14.74) | (4.239) |
| 2003-2007 | 1,710\*\*\* | 1,567\*\*\* | 1,602\*\* | 68.58\*\*\* | 61.14\*\*\* | 65.83\*\*\* |
|  | (537.3) | (548.2) | (629.9) | (12.02) | (22.47) | (11.48) |
|  |  |  |  |  |  |  |
| **面板K: 保留没有进行跨城市流动的本地样本** | | | | | | |
| 2002 | 286.0 | 88.12 | 362.3 | 10.22\* | 2.837 | 11.11\*\*\* |
|  | (241.2) | (284.7) | (228.6) | (5.246) | (8.407) | (3.471) |
| 2003-2007 | 1,609\*\*\* | 1,379\*\*\* | 1,622\*\*\* | 64.55\*\*\* | 55.81\*\*\* | 64.35\*\*\* |
|  | (493.7) | (475.7) | (573.9) | (7.621) | (15.74) | (10.54) |
|  |  |  |  |  |  |  |
| 城市和年份固定效应 | 是 | 是 | 是 | 是 | 是 | 是 |
| 性别、年龄及交互项 | 是 | 是 | 是 | 是 | 是 | 是 |

注：面板A数据来源为UHS 1996-2007年样本，其他面板数据来源为UHS 1999-2007年样本。其中第（1）-（3）列中的处理变量为根据城市1998年高等院校数构建得到的二元分组变量，第（4）-（6）列中的处理变量为根据城市1998年高等院校数构建得到的连续强度变量。括号内是聚类在城市层面的稳健性标准误，\*、\*\*、\*\*\*分别表示在10%、5%、1%的水平上显著。

附录V 人力资本外溢效应的潜在机制：非参数动态双重差分估计结果

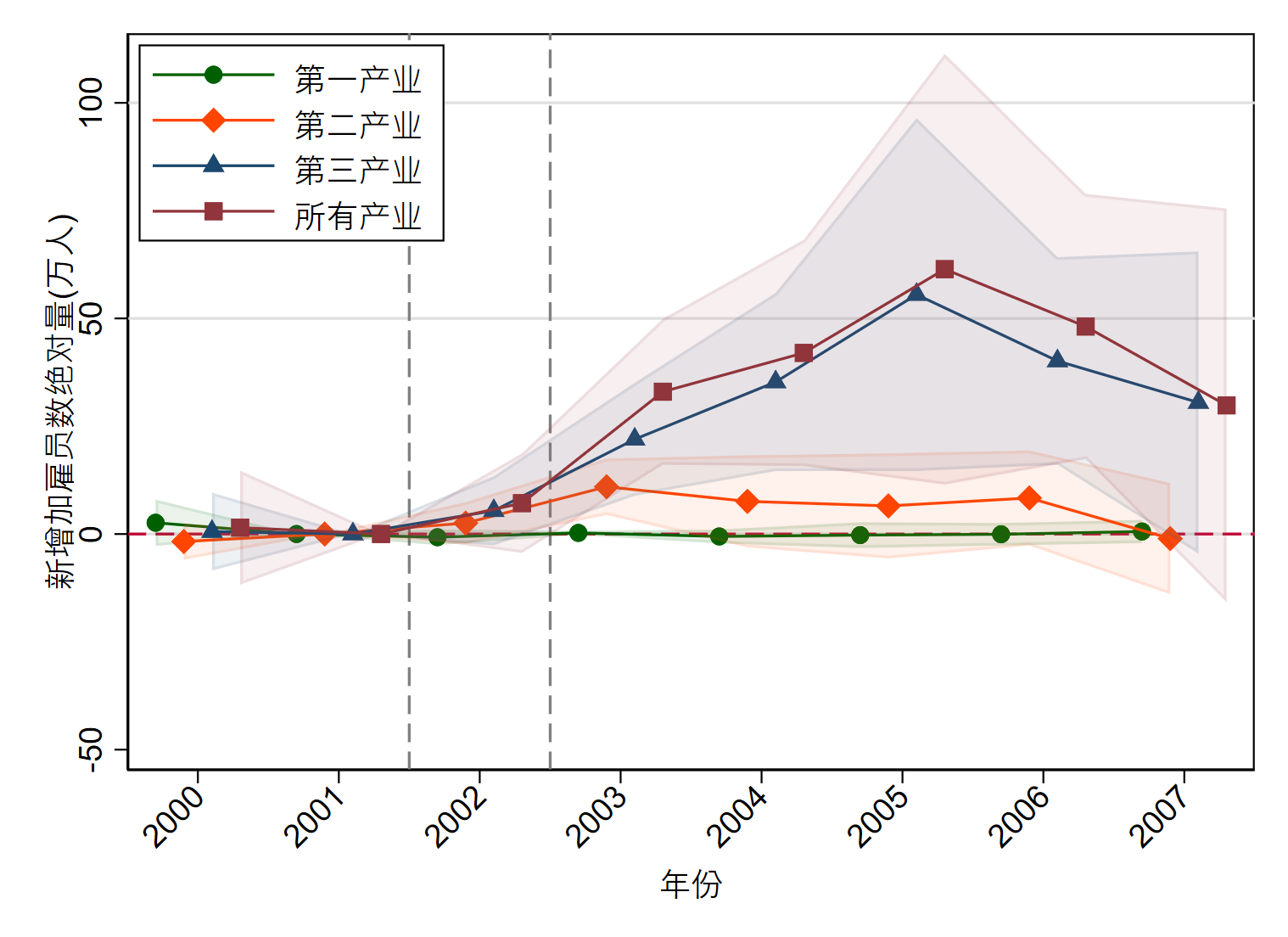


图 V1 扩招政策对城市层面劳动需求的影响：非参数动态双重差分估计结果

注：数据来源为工商注册企业数据2000-2007年样本。非参数动态双重差分估计结果为根据正文模型（2）进行估计得到，阴影区域为系数的95%置信区间。

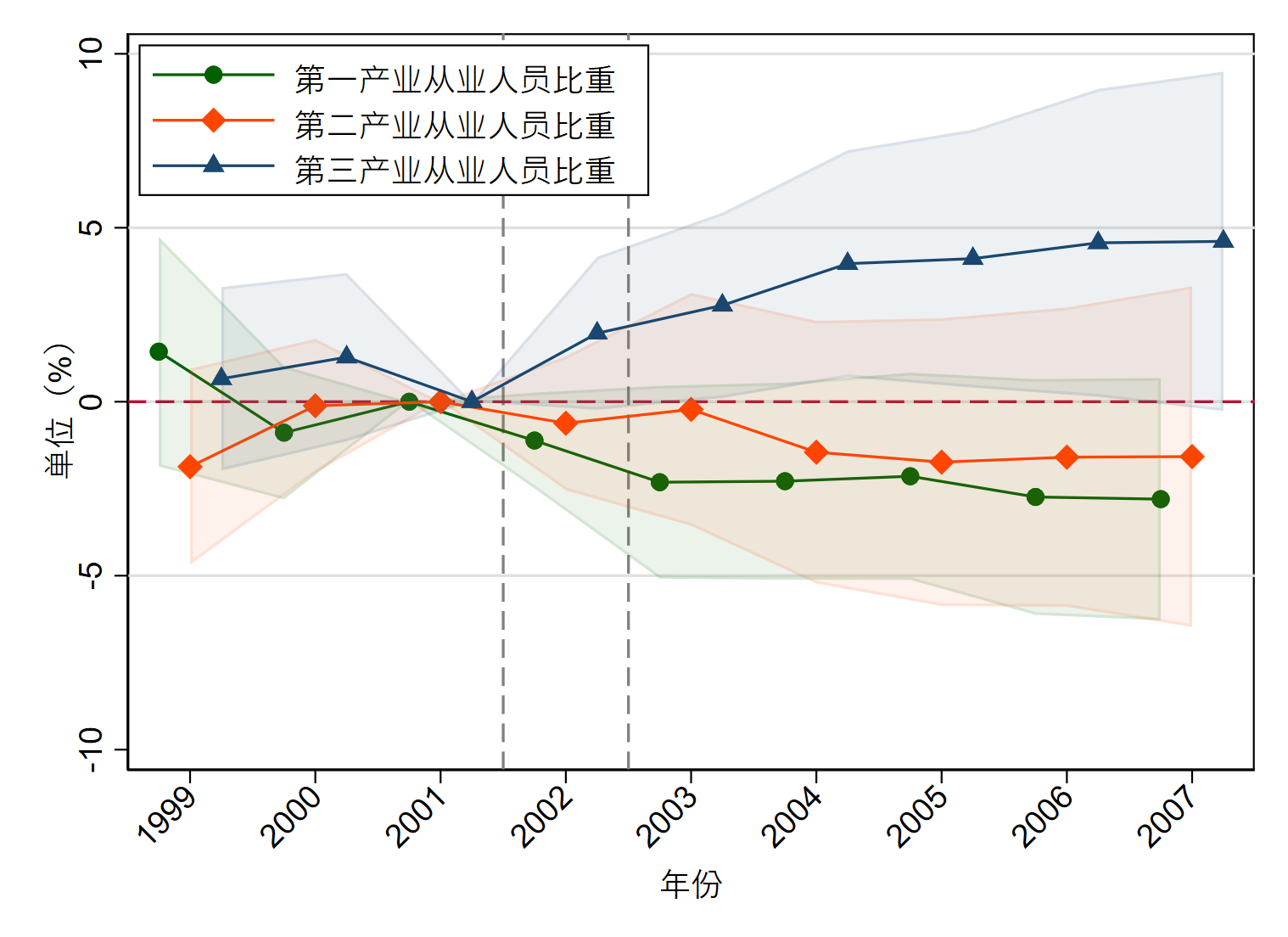
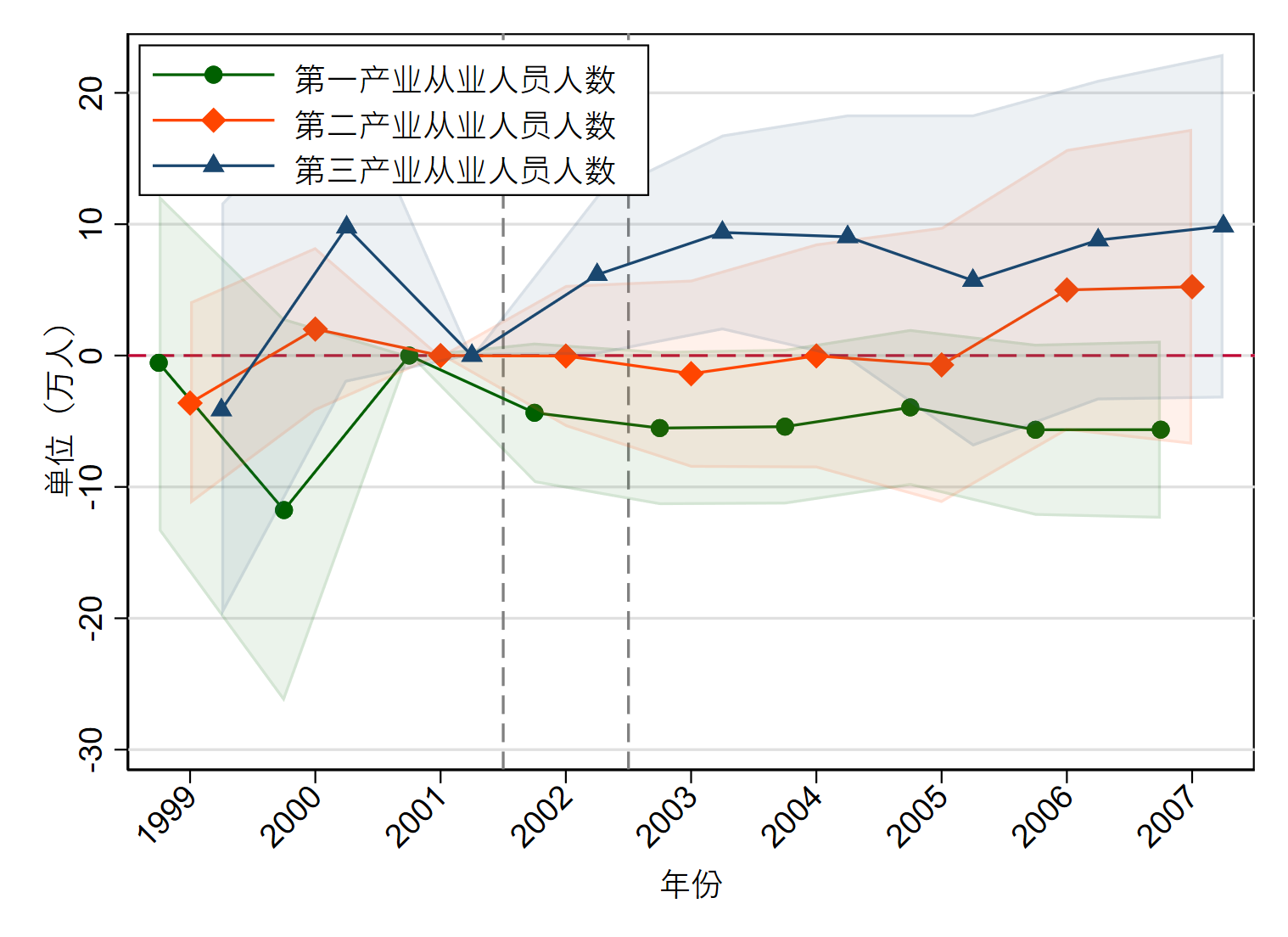


图 V2 扩招政策对城市产业结构的影响：非参数动态双重差分估计结果

注：数据来源为城市统计年鉴数据1999-2007年样本。非参数动态双重差分估计结果为根据正文模型（2）进行估计得到，阴影区域为系数的95%置信区间。

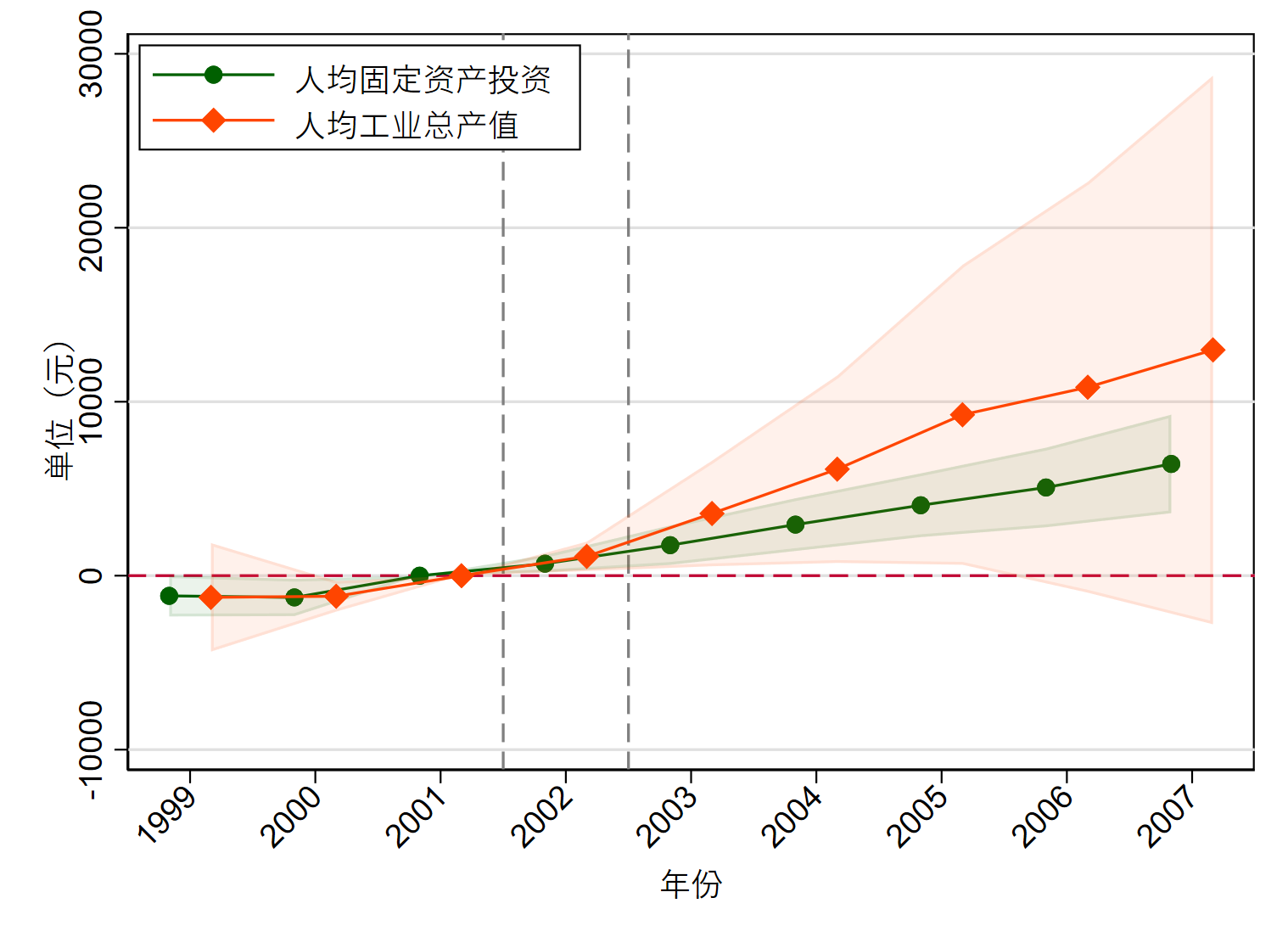


图 V3 扩招政策对城市工业发展的影响：非参数动态双重差分估计结果

注：数据来源为城市统计年鉴数据1999-2007年样本。非参数动态双重差分估计结果为根据正文模型（2）进行估计得到，阴影区域为系数的95%置信区间。

附录VI 关于大学教育回报的进一步讨论：分析框架和估计结果

**表 VI1 扩招政策外溢效应的分析框架**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 处理组城市相对控制组城市 | 无外溢效应 | | 存在外溢效应 | |
|  | 扩招前 | 扩招后 | 扩招前 | 扩招后 |
| 受政策直接影响群体 |  |  |  |  |
| 未受政策直接影响群体 |  |  |  |  |
| OLS估计结果 |  | |  | |
| 平均处理效应 |  | |  | |

注：在本文中，经典双重差分模型的两个差分维度分别为年份层面（根据2003年前后划分）和城市层面（根据1998年城市高校数量划分），因此表中、均为城市层面的组间差异（收入差分后结果），前者表示2003年之前，处理组城市与控制组城市中受政策直接影响群体的平均收入之差，后者表示2003年之前，处理组城市与控制组城市中未受政策直接影响群体的平均收入之差。此外，由于表示每个大学生带来的对自身和其他人收入改善的总效果，因此所有受政策直接影响群体带来的溢出效应效果等价于政策实施时未完成教育群体的额外人均收入增加与政策实施时已完成教育群体的额外人均收入增加之和。

我们基于Wald统计量的概念，分别使用两类“工具变量”进行2SLS估计，一类为二元分组变量（根据城市1998年高等院校数构建得到）与2003年前后的交互项，相应结果记为2SLS-1并报告在表V2的第（2）列和第（5）列中；另一类为连续强度变量（城市1998年高等院校数）与2003年前后的交互项，相应结果记为2SLS-2并报告在第（3）列和第（6）列中 。结果显示，在全样本中，每个大学生每年带来的人力资本外溢效应约为56,000~76,000元，且扩招政策带来的人力资本个人回报均明显低于社会回报（约为后者的9%~13%），与正文结论基本保持一致，进一步证实了人力资本外溢效应的存在和教育回报的低估问题。此外，在出生于1960年后的子样本中，我们也发现了类似结论，证实了结论的稳健性。

**表 VI2 人力资本个人回报和社会回报的估计结果**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | | (3) | | (4) | | (5) | | (6) | |
| 变量 | 个人收入 （元） | | | | | | | | | | |
| 样本 | 全样本 | | | | | 生于1960年后 | | | | | |
| 估计方法 | OLS | | 2SLS-1 | | 2SLS-2 | | OLS | | 2SLS-1 | | 2SLS-2 |  |
|  | | | |  | |  | |  | |  | |
| 上过大学 | 6,507\*\*\* | 56,837\*\*\* | | 76,747\*\*\* | | 5,627\*\*\* | | 30,192\*\*\* | | 38,639\*\*\* | |  |
| （是=1） | (439.4) | (19,251) | | (20,452) | | (396.0) | | (8,678) | | (8,832) | |  |
|  |  |  | |  | |  | |  | |  | | |
| Hansen | -- | 1.304 | | 1.977 | | -- | | 3.593 | | 3.426 | | |
| P值 | -- | 0.861 | | 0.740 | | -- | | 0.464 | | 0.489 | | |
|  |  |  | |  | |  | |  | |  | |
| 城市和年份固定效应 | 是 | 是 | | 是 | | 是 | | 是 | | 是 | |
| 性别、年龄及交互项 | 是 | 是 | | 是 | | 是 | | 是 | | 是 | |
| 样本量 | 275,998 | 275,998 | | 275,998 | | 156,765 | | 156,765 | | 156,765 | |

注：数据来源为UHS 1999-2007年样本，相关结果为根据OLS和2SLS方法进行估计得到，其中2SLS-1基于城市1998年高等院校数的二元分组变量与2003年前后的交互项进行估计，2SLS-2基于连续强度变量与2003年前后的交互项进行估计。括号内是聚类在城市层面的稳健性标准误，\*、\*\*、\*\*\*分别表示在10%、5%、1%的水平上显著。

**参考文献**

1. 葛润、黄家林，“大学扩招是否影响了结婚与生育”，《经济学报》，2020年第3期，第168—201页。
2. 宁光杰、段乐乐，“流动人口的创业选择与收入——户籍的作用及改革启示” ，《经济学（季刊）》， 2017年第2期，第771—792页。
3. 张原、陈建奇，“人力资本还是行业特征:中国行业间工资回报差异的成因分析”，《世界经济》，2008年第5期，第68—80页。
4. 周茂、李雨浓、姚星、陆毅，“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》，2019年第5期，第64—77+198—199页。

**注：该附录是期刊所发表论文的组成部分，同样视为作者公开发表的内容。如研究中使用该附录中的内容，请务必在研究成果上注明附录下载出处**。

1. 在1999年推动的高校扩招等大学教育改革中，其中的高等院校既包括大学等本科院校，也包括职业学院等专科院校。因此，为了尽可能准确地捕捉政策影响，我们将受教育程度为大学专科及以上的群体均定义为大学毕业生。 [↑](#footnote-ref-0)
2. 个人收入变量为包含工薪收入、经营净收入、财产性收入、转移性收入等的总收入，且该变量是在年份层面进行测度。我们选择个人总收入的原因在于该指标能够较为全面地反映个体收入的变化，进而保证人力资本外溢效应效果估计的准确性。 [↑](#footnote-ref-1)
3. 例如，对于一个子代出生于1960年后、父代出生于1960年前的家庭而言，尽管父代不会因直接受到扩招政策影响而获得人力资本提高以及个人收入改善，但由于子代在获得人力资本提高后可以通过家庭内部的人力资本溢出效应途径改善父代的收入状况，因此最终可以观察到，扩招政策使得出生于1960年前队列中处理组的个人收入也出现了显著提高。 [↑](#footnote-ref-2)