**家政服务业发展提升了女性就业水平**

梁超 熊瑞祥 王素素 孙贺贺

**目录**

[附录I 家政服务体系建设试点城市名单 1](#_Toc15616424)

[附录II CHNS数据样本筛选过程 2](#_Toc472233234)

[附录III CHNS数据描述性统计 3](#_Toc1871888173)

[附录IV 基于普查数据考察家政服务体系建设对女性就业的影响 4](#_Toc189095061)

[附录V 稳健性检验 5](#_Toc1992376314)

[附录VI 基于CHNS数据考察家政服务体系建设对高低学历女性的异质影响 10](#_Toc156201727)

[附录VII基于UHS数据考察家政服务体系建设对家政服务消费的影响 11](#_Toc1057409055)

附录I 家政服务体系建设试点城市名单

截至2012年，总共有91个城市先后入选家政服务体系建设试点名单。具体试点城市附表Ⅰ1所示。其中加粗城市为CHNS覆盖的试点城市。

表I1 家政服务体系建设试点城市名单

|  |  |
| --- | --- |
| 2010 | 北京市、天津市、廊坊市、秦皇岛市、太原市、呼和浩特市、大连市、吉林市、**哈尔滨市、**上海市 、**南京市、**杭州市、合肥市、福州市、厦门市、南昌市、潍坊市、**郑州市、武汉市、长沙市、**郴州市、广州市、深圳市、**南宁市、**成都市、万州市、**贵阳市、**拉萨市、昆明市、西安市、延安市、兰州市、西宁市、银川市、石河子市 |
| 2011 | 邢台市、晋中市、运城市、包头市、锦州市、辽源市、绥化市、南通市、金华市、亳州市、泉州市、赣州市、**济南市、**威海市、周口市、咸宁市、孝感市、湘潭市、张家界市、惠州市、中山市、桂林市、三亚市、德阳市、绵阳市、遵义市、曲靖市、日喀则市、汉中市、酒泉市、白银市、固原市、格尔木市、克拉玛依市、重庆市、宁波市、青岛市 |
| 2012 | 菏泽市、铜陵市、蚌埠市、玉林市、石家庄市、邯郸市、襄阳市、绍兴市、三明市、宁德市、榆林市、大庆市、海口市、儋州市、晋城市、长治市、**十堰市**、库尔勒市、淮安市 |

数据来源：根据商务部以及各省商务厅文件手动整理。

附录II CHNS数据获取及样本筛选过程

由于CHNS数据并未提供城市和区县代码，本文参考Chyi & Zhou（2014）的方法进行识别。尽管CHNS数据未提供城市和区县代码，但是在社区数据模块中提供了所在县（市）的人口规模和面积信息。综合多本统计年鉴资料获取的县（市）信息，包括《中国县(市)社会经济统计年鉴》《中国城市统计年鉴》《中国县域统计年鉴》《广西统计年鉴》《黑龙江统计年鉴》和《江苏统计年鉴》，上述材料中提取出对应调查年份的县（市）人口规模和面积信息。根据省份ID结合人口规模、面积信息进行匹配，在程序识别尽可能减少误差的基础上辅以人工二次识别和双重校准，最终获得CHNS调查村居（社区）所对应的48个城市、56个县（市）信息，并根据地级市特征与本文的政策实施信息进行匹配。事实上也有其他基于CHNS区县代码展开的研究采用了该方法进行识别，比如Fang & Cao（2020）。

本文对所有女性样本都限定了22-55岁范围，具体筛选过程表Ⅱ1所示：（1）为避免过长的时期引入过多的干扰因素，本文仅使用2006、2009、2011和2015四个批次的调查，样本由初始的110120减少到40088；（2）由于家政服务市场属于面对面服务业，其诞生和发展都需要一定的人口密度，因此本文限制使用城镇地区样本，此时样本数量进一步减少到13266；（3）本文关注家政服务业发展对女性就业的影响，仅保留女性使得样本减少到6954；（4）既然是研究女性就业问题，那么只关注劳动年龄人口女性，限制年龄范围为22-55岁，样本进一步减少到3931。以上这就是本文基础回归所用样本的筛选过程。

表II1 CHNS数据样本选择过程

|  |  |  |
| --- | --- | --- |
| 处理步骤 | 丢掉样本 | 保留样本 |
| CHNS原始样本 |  | 110120 |
| 保留2006、2009、2011、2015年调查数据 | 70032 | 40088 |
| 保留城市样本 | 26822 | 13266 |
| 保留女性样本 | 6312 | 6954 |
| 限制年龄范围22-55 | 3023 | 3931 |
| 最终样本 |  | 3931 |

附录III CHNS数据描述性统计

表III1显示，整体而言劳动年龄女性的就业率仅为58.6%，仍有较大的上升空间。表III2进一步将样本分成家政服务体系建设试点城市与非试点城市，以及试点前与试点后。表III2显示，在试点后，试点城市的女性就业参与率增加了约3.07%、每周工作时长增加了约2.2小时；相比之下，非试点城市的女性样本的就业参与率和周工作时长没有明显的增加，二者分别减少了约0.216%与0.238小时。表III2的描述统计初步表明，家政服务体系建设政策提高了试点地区的女性就业概率与工作时长。此外，还可以发现非试点地区女性照料孩子时长在试点前后增加2.2个小时，但该变化并不显著。同样的，试点地区女性照料孩子时长在试点前后减少1.88个小时，变化同样不显著。

表III1 描述性统计-全样本

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 样本量 | 均值 | 标准差 | 最小值 | 最大值 |
| 就业参与 | 3931 | 0.586 | 0.493 | 0 | 1 |
| 周工作时长(小时) | 3806 | 21.832 | 22.074 | 0 | 100 |
| 做家务时长(小时/周) | 3773 | 17.257 | 12.437 | 0 | 112 |
| 照顾老人时长(小时/周) | 3931 | 2.461 | 11.703 | 0 | 140 |
| 照顾孩子时间(小时/周) | 3784 | 5.044 | 16.959 | 0 | 168 |
| 年龄 | 3931 | 41.429 | 9.293 | 22 | 55 |
| 教育年限 | 3925 | 9.946 | 3.758 | 0 | 18 |
| 在婚 | 3926 | 0.862 | 0.345 | 0 | 1 |
| 农村户口 | 3901 | 0.355 | 0.478 | 0 | 1 |
| 家庭成员数量 | 3801 | 3.655 | 1.337 | 1 | 10 |

资料来源：根据2006、2009、2011、2015年CHNS数据整理。

表III2 描述性统计-试点城市与非试点城市在试点前后的差异

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 试点城市样本 | 试点前 | | | 试点后 | | |  |
|  | 样本量 | 均值 | 标准差 | 样本量 | 均值 | 标准差 | 变化 |
| 就业参与 | 979 | 0.563 | 0.496 | 898 | 0.594 | 0.491 | 0.0307 |
| 周工作时长(小时) | 955 | 20.155 | 21.338 | 864 | 22.356 | 22.415 | 2.200\*\* |
| 做家务时长(小时/周) | 930 | 18.406 | 12.959 | 874 | 14.776 | 12.084 | -3.630\*\*\* |
| 父母公婆是否需要照顾 | 979 | .152 | .359 | 898 | .125 | .331 | -0.0275\* |
| 照顾老人时长(小时/周) | 149 | 6.725 | 14.569 | 112 | 7.75 | 16.88 | 1.025 |
| 是否有六岁以下儿童 | 979 | .158 | .365 | 898 | .198 | .399 | 0.0399\*\* |
| 照顾孩子时间(小时/周) | 145 | 25.731 | 30.547 | 172 | 23.849 | 29.015 | -1.882 |
| 非试点城市样本 | 试点前 | | | 试点后 | | |  |
|  | 样本量 | 均值 | 标准差 | 样本量 | 均值 | 标准差 | 变化 |
| 就业参与 | 1149 | 0.594 | 0.491 | 905 | 0.592 | 0.492 | -0.00216 |
| 周工作时长(小时) | 1101 | 22.535 | 22.559 | 886 | 22.297 | 22.011 | -0.238 |
| 做家务时长(小时/周) | 1078 | 19.424 | 12.457 | 891 | 15.867 | 11.583 | -3.557\*\*\* |
| 父母公婆是否需要照顾 | 1149 | .149 | .356 | 905 | .151 | .359 | 0.0026 |
| 照顾老人时长(小时/周) | 171 | 8.14 | 20.32 | 137 | 9.81 | 23.174 | 1.670 |
| 是否有六岁以下儿童 | 1149 | .195 | .396 | 905 | .149 | .356 | -0.458\*\*\* |
| 照顾孩子时间(小时/周) | 196 | 24.133 | 31.729 | 131 | 26.405 | 27.546 | 2.272 |

数据来源：根据2006、2009、2011、2015年CHNS数据整理。

**附录**IV 基于普查数据考察家政服务体系建设对女性就业的影响

根据审稿专家的意见，这里尝试基于2010和2015年人口普查数据讨论家政服务体系建设对女性工作参与的影响。尽管2010年普查是在11月份进行，但是这里仍然将2010当年视为政策处理前，将2010至2012年所有实施过家政服务体系建设试点的城市视为处理组。基于公式（2）的双重差分模型进行回归，控制变量包括女性的年龄及其平方项、教育年限、婚姻状态、上一年是否生育子女等信息，在控制城市与年份固定效应的基础上，也进一步控制住城市初始特征与年份的交叉项，以减弱潜在的政策选择内生性问题。回归结果如表Ⅳ1所示。列1和列2显示，家政服务体系建设带来整体上女性就业参与率提高2.3%到2.5%。列3和列4显示，家政服务体系建设对初中及以下学历女性的就业参与率提高作用更为明显，使得她们的就业率提高3.37%，对于高中及以上学历女性的就业参与率提高1.37%。这里的考察与基于CHNS数据的考察结果基本一致，但是经济显著性有较大的差异，可能的原因在于：首先是时间范围差异，这里仅使用了2010和2015年的数据，而CHNS使用了2006年以来的四次调查数据；其次是处理组设定，这里设定2010年受处理城市的2010年视为政策处理前，对于政策效应也会带来一定程度的低估。

表IV1 基于普查数据考察家政服务体系建设对女性就业参与的影响

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | 女性样本 | 女性样本 | 初中及以下学历女性 | 高中及以上  学历女性 |
| 变量名称 | 是否就业 | 是否就业 | 是否就业 | 是否就业 |
| 家政服务体系建设 | 0.0232\*\* | 0.0249\*\*\* | 0.0337\*\*\* | 0.0137\* |
|  | (0.00910) | (0.00891) | (0.0116) | (0.00814) |
| 样本量 | 377,862 | 377,862 | 159,991 | 217,871 |
| 调整R平方 | 0.097 | 0.097 | 0.059 | 0.091 |
| 城市数量 | 281 | 281 | 281 | 281 |
| 控制变量 | 控制 | 控制 | 控制 | 控制 |
| 城市FE | 控制 | 控制 | 控制 | 控制 |
| 年份FE | 控制 | 控制 | 控制 | 控制 |
| 初期特征\*年份FE |  | 控制 | 控制 | 控制 |

注: 括号中是聚类到城市层面的稳健标准误，\*\*\*、\*\*、\*分别表示 1% 、5% 、10%的显著水平。个人与家庭特征包括：年龄及其平方项、是否在婚、是否农村户口、家庭成员数量。城市层面的期初特征包括：2009年人均GDP对数、2010年14岁以下少儿和65岁以上老年人占比、劳动年龄女性就业率。

附录V 稳健性检验

**1. 处理交错型双重差分（Staggered DID）估计中的潜在偏误**

当使用双向固定效应方法估计渐进DID（交错DID）模型时，往往存在异质性处理效应和负权重的问题，造成估计结果偏误（De Chaisemartin & D'Haultfoeuille, 2020a；Callaway & Sant’ Anna, 2021）。针对交错型DID模型，Goodman-Bacon（2021）根据受政策冲击时间将样本分为早处理组、晚处理组和从未接受处理组，通过DID分解研究了偏误产生的来源，并进一步提出通过加权方差的共同趋势来判断估计量偏误程度的方法。但是Goodman-Bacon（2021）的方法要求数据为完全平衡面板，本文的数据结构难以满足该要求。本文中的18个样本城市中有9个实施了家政服务体系建设试点，其中有7个属于2010年公布的第一批试点名单，有2个分别属于2011和2012年公布的试点名单。为处理交错型DID可能出现的估计偏误问题，这里仅保留2010年第一批试点城市样本和一直未实施试点的对照组样本进行估计，此时的DID回归属于经典的2\*2形式，不存在异质性处理效应和负权重问题，且样本损失较少。此时避免了使用较早处理组（即“坏的控制组”）作为较晚处理组的对照组，而这恰好是双向固定效应产生估计偏误的核心问题，由此避免了估计偏误。事实上，这里的处理方法与Sun & Abraham（2021）针对异质性处理效应提出的IW估计量的计算过程有类似之处，后者在遇到不存在一直未处理组（Never-Treated Group）时，采用删除最后处理组样本的方式来构造合理的控制组。表Ⅴ1报告相应的估计结果：家政服务体系建设试点政策仍然显著地提高了女性的就业概率与周工作时长，进一步表明前文估计结果是可靠的。

表V1 处理交错型双重差分估计中的潜在偏误

|  |  |  |
| --- | --- | --- |
|  | (1) | (2) |
| 变量名称 | 是否就业 | 周工作小时数 |
| 家政服务体系建设 | 0.0819\* | 4.577\*\* |
|  | (0.0438) | (2.113) |
| 样本量 | 3,409 | 3,319 |
| 调整R平方 | 0.459 | 0.401 |
| 个人与家庭特征 | 控制 | 控制 |
| 个体FE | 控制 | 控制 |
| 年份FE | 控制 | 控制 |
| 期初特征\*年份FE | 控制 | 控制 |

注: 括号中为bootstrap稳健标准误（抽样次数为1000），\*\*\*、\*\*、\*分别表示 1% 、5% 、10%的显著水平。个人与家庭特征包括：年龄及其平方项、受教育年限、是否在婚、是否农村户口、家庭成员数量。城市层面的期初特征包括：2009年人均GDP对数、2010年14岁以下少儿和65岁以上老年人占比、劳动年龄女性就业率。

这里剔除两个城市的做法有可能带来样本选择（sample selection）的问题[[1]](#footnote-0)。作为稳健性检验，我们也采用Callaway & Sant'Anna (2021)提出的依赖计算组别-时期平均处理效应的方法，以及Borusyak et al.(2021)和Gardner (2021)提出的采用插补估计量构造反事实估计的方法考察家政服务体系建设试点政策的平均处理效应。之所以未采用De Chaisemartin & D'Haultfeuille（2020b）和Sun & Abraham(2021)的方法，是因为前者在估计平均处理效应时，仍然要求给出至少一期的动态效应，后者则需要先估计event study结果，然后对其进行加权求均值得到平均处理效应。具体估计结果如表V2所示。可以发现，采用三种异质性稳健估计量仍然能否发现家政服务体系建设试点政策对女性劳动参与和周工作时长的正面效应，不过由于估计中采用的样本量差异，估计系数大小有所变化，但均显著为正。

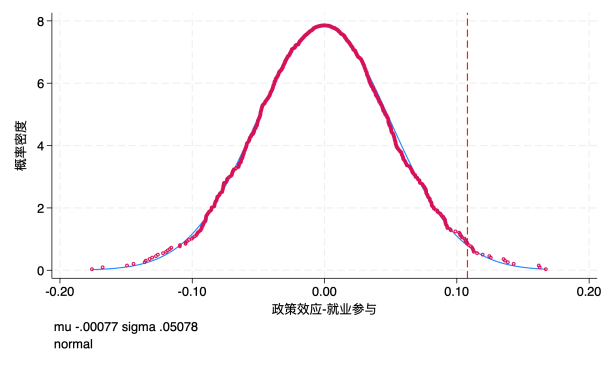
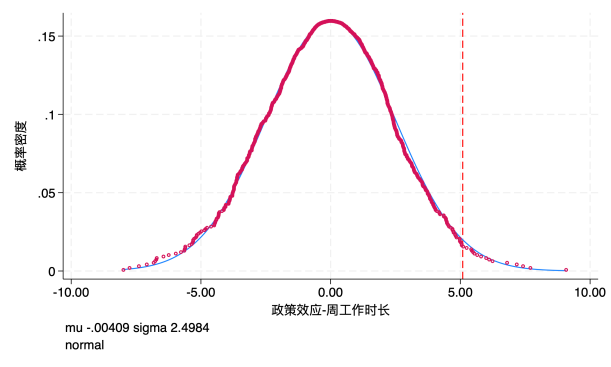
表V2 考虑交错型双重差分估计偏误的稳健性检验

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 方法 | Callaway & Sant'Anna (2021) | | Borusyak et al.(2021) | | Gardner (2021) | |
|  | 就业与否 | 周工作时长（小时） | 就业与否 | 周工作时长（小时） | 就业与否 | 周工作时长（小时） |
| 家政服务体系建设试点 | 0.0685\*\* | 2.743\* | 0.0590\* | 3.161\*\* | 0.0504\*\* | 1.921\*\* |
| (0.0349) | (1.598) | (0.0317) | (1.459) | (0.0205) | (0.931) |
| 样本量 | 1547 | 1547 | 3527 | 3527 | 1547 | 1547 |

注：Callaway & Sant'Anna (2021)和Gardner (2021)的方法依赖严格的平衡面板数据，此时样本量减少到1547。

**2. 随机生成试点城市进行安慰剂检验**

为进一步检验本文的估计结果是否是由其他不可观测因素驱动，本小节参考Chetty et al.(2009)以及Li et al.(2016)的做法，通过随机生成试点城市与执行年份来进行安慰剂检验。本文18个城市样本中有9个城市先后进入处理组，9个试点城市在三个试点年份上的分布分别为7、1、1，这里采用如下步骤构造安慰剂检验：第一步，随机选择一个年份作为试点年份，在18个城市中随机选择7个城市作为试点城市；第二步，将第一步选出的试点年份和试点城市剔除，在剩下年份中选择一个试点年份，然后选择一个试点城市；第三步，在剔除前两步选中的城市和年份后，剩下的年份中选择一个试点年份，然后选择一个试点城市；第四步，将虚拟选出的试点城市-年份数据与微观数据进行匹配，根据调查年份和虚拟试点城市、年份构造出虚拟的双重差分处理变量信息，进行回归得到虚拟家政服务体系建设试点政策的系数。对女性就业概率与周工作小时数这两个被解释变量，这里各进行了1000次随机抽样，并根据回归方程（2）式进行了1000次回归。因为这些处理城市与处理时间是随机选取的，故理论上，平均而言，虚拟的政策变量对女性就业应该没有显著影响。图V1与V2的安慰剂检验结果分别报告了被解释变量为女性就业参与、女性周工作小时数时，关键的虚拟解释变量估计系数的分布。如图所示，家政服务体系建设试点对女性就业参与和周工作时长的促进效应都不是随机得到的。

图V1 虚拟政策对女性是否就业的影响 图V2 虚拟政策对女性周工作小时数的影响

**3. 控制省级层面随时间变化的不可观测因素**

省级层面随时间变化的不可观测因素，也可能既同关键解释变量相关，又影响被解释变

量，这些不可观测因素可能干扰估计结果。例如，省级层面的经济波动可能既影响各省家政服务体系建设试点政策的实施，也可能通过劳动力市场需求影响女性的就业行为与工作时长。为此，在基准回归基础上，这里进一步控制省份固定效应与年份固定效应的交互项，表V3第1行的估计结果显示，此时家政服务体系建设试点政策仍然显著地增加了女性的就业概率与每周工作小时数，其中解释变量为周工作小时数时关键解释变量系数的P值为0.102。

**4. 排除城市经济增长、家庭收入增长的影响**

前文发现的女性就业概率增加和工作时间延长也可能受其他诸如收入增长等需求因素的影响，而非受家政服务体系建设的影响。例如，城市经济增长可能带来新的就业机会，从而吸引女性进入劳动力市场和增加劳动时长；家庭收入的上涨也会使得家庭更有能力购买家政服务，从而缓解女性的家务劳动与家庭照料负担，促进其就业。为此，在基准回归的基础上，这里进一步控制住城市人均GDP的对数与经过CPI调整后的家庭人均收入的对数，[[2]](#footnote-1) 表V3第2行的估计结果显示，此时家政服务体系建设试点政策仍然显著地增加了女性的就业概率与每周工作小时数。

**5. 排除城市服务业发展的影响**

服务业发展本身有助于促进女性就业。如果家政服务体系建设试点城市的服务业发展潜力越大，服务业发展的水平越高，女性的就业概率也会更高、工作时间也会更长。为此，在基准回归的基础上，这里控制城市初期的服务业占GDP比例与年份固定效应的交互项。表V3第3行的估计结果显示，此时家政服务体系建设仍然显著地增加了女性的就业概率与每周工作小时数。进一步控制住城市层面历年服务业占GDP比例，表V3第4行的估计结果显示，此时家政服务体系建设试点政策仍然显著地促进了女性就业水平提高。本部分考察说明并非城市服务业整体发展水平的提高驱动了前文的基准结果。

**6. 控制女性地位的影响**

传统性别角色也会影响女性的就业参与：性别角色观念越传统的地区，女性从事受雇工作的概率越低（张川川和王靖雯，2020；续继和黄娅娜，2018）。地区的性别传统角色也可能影响当地的家政服务体系建设，因此，忽视城市层面的女性地位因素就可能影响估计结果。本文以城市婚姻市场上适龄人群性别比来衡量当地的女性家庭地位，设定适婚人群的年龄范围为22-35岁。基于2010年人口普查数据采用群组推算的方式得到2006、2009、2011和2015年样本城市的婚龄人群性别比：例如，以2010年某地级市23-36岁人群的性别比来代表2009年该地级市22-35岁婚恋人群性别比。前述基准回归基础上，这里控制住城市层面历年婚龄人群的性别比，表V3第5行的估计结果显示，此时家政服务体系建设试点政策对女性就业水平的促进效应依然显著。

**7. 排除幼儿与老人照料公共服务供给的影响**

小学以及幼儿园等公共服务的供给增加可以缓解照顾儿童对女性非农就业的阻碍效应（熊瑞祥和李辉文，2016），养老院的增加也有助于缓解照料老年人对女性就业的负面影响（范红丽和辛宝英，2019）。近些年我国政府大力推进基本公共服务均等化，例如，修建更多的幼儿园和养老院。如果试点城市在推进家政服务业发展的同时新建了更多的幼儿园和养老院，也可以降低女性日常照料子女和老人的负担，促进女性就业并延长她们的周工作时长。这意味着，前文家政服务体系建设的效应可能捕捉了其他公共服务设施扩张的影响。为此，在基准回归的基础上，这里控制住城市层面历年每万常住人口幼儿园数量，表V3第6行的估计结果显示，此时家政服务体系建设试点政策仍然显著地增加了女性的就业概率与每周工作小时数。此外，本部分还控制住每万常住人口养老床位数与年份哑变量的交互项，[[3]](#footnote-2)排除养老资源扩张可能的影响，此时家政服务体系建设试点政策对女性就业的促进效应依然稳健。

表V3 各类稳健性检验

|  |  |  |
| --- | --- | --- |
|  | (1) | (2) |
| 稳健性检验方式 | 是否就业 | 周工作小时数 |
| 省份固定效应\*  年份固定效应 | 0.242\*\* | 8.245\* |
| (0.104) | (4.885) |
| 城市人均GDP与  家庭人均收入 | 0.108\*\*\* | 4.947\*\* |
| (0.0402) | (1.979) |
| 城市初期服务业占GDP比例\*  年份固定效应 | 0.147\*\* | 6.528\*\* |
| (0.0582) | (2.873) |
| 城市历年服务业占GDP比例 | 0.0991\*\* | 4.556\*\* |
|  | (0.0410) | (2.009) |
| 城市历年婚龄人群性别比 | 0.0991\*\* | 4.849\*\* |
| (0.0430) | (2.102) |
| 城市历年每万人幼儿园数量 | 0.109\*\*\* | 5.065\*\* |
| (0.0411) | (1.998) |
| 城市每万人养老院床位数\*  年份固定效应 | 0.121\*\*\* | 5.297\*\* |
| (0.0444) | (2.194) |
| 以标准年龄段人口考察（16-55岁女性） | 0.106\*\*\* | 4.987\*\*\* |
| (0.0382) | (1.916) |
| 排除移民的潜在干扰 | 0.122\*\* | 6.276\*\* |
| (0.0533) | (2.536) |
| 剔除观测样本量较少的城市 | 0.124\*\*\* | 5.739\*\*\* |
| (0.0424) | (2.039) |

注: 括号中为bootstrap稳健标准误（抽样次数为1000），\*\*\*、\*\*、\*分别表示 1% 、5% 、10%的显著水平。所有回归中控制住个体和年份固定效应、个体与家庭特征以及城市层面期初特征与年份哑变量的交叉项。个人与家庭特征包括：年龄及其平方项、受教育年限、是否在婚、是否农村户口、家庭成员数量。城市层面的期初特征包括：2009年人均GDP对数、2010年14岁以下少儿和65岁以上老年人占比、劳动年龄女性就业率。

**8. 采用标准劳动年龄**

本部分尝试以标准的16-55岁筛选劳动年龄女性作为稳健性检验，此时样本量有所扩大。结果如表V3第8行所示，可以发现即使采用标准的劳动年龄女性样本，家政服务体系建设对女性的就业参与和周工作时长的正面影响依然稳健。需要说明的是，之所以未将16-55岁女性人群的考察作为基准回归，是因为随着中国高等教育的普及率不断升高，超过16岁仍在求学的比例越来越高，本文2015年CHNS样本中有超过66%的16-21岁女性仍然处在上学阶段。

**9. 排除移民的潜在干扰**

本部分讨论移民的存在可能造成的影响。如果家政服务体系建设试点政策的确更多促进了低学历女性的就业，同时移民也以低技能人群为主，且多因为工作而迁移的话，那么移民的存在可能造成对女性就业系数的低估。这里尝试将移民信息纳入，以排除移民进行考察。[[4]](#footnote-3)遗憾的是，CHNS数据中仅在2015年问到受访者“是否出生在本省”以及“出生地是城市、郊区、县城还是农村”，根据出生地信息以及当前受访所在地区是否城区，我们将未出生在本地的以及当前所在地是城区但出生地是农村的样本定义为移民（严格来讲这里既包含了跨地区移民也包含了本地的城乡间移民），并根据2015年个体id向前匹配，仅能得到部分样本的移民信息（因为有一些观测样本未在2015年出现）。仅使用本地非移民样本的考察结果如表V3第9行所示。可以发现，家政服务体系建设试点对女性的劳动参与和周工作时间的促进效应依然显著。值得关注的是，排除移民后的效应大小超过包含移民时的影响，说明移民的存在使得基准考察效应存在一定程度的低估。考虑到本文样本量较少，而且低技能劳动力中有相当比例的移民人群，为了综合反映家政服务业发展的影响，在基准回归中仍将移民纳入考察范围，采用控制个体固定效应的方式缓解移民的潜在影响。

**10. 排除抽样样本量较少城市的干扰**

由于本文所用CHNS数据观测样本量较少，为避免一些城市样本量过少带来的潜在干扰，这里将样本量少于30的城市予以剔除。这样能保证留下的城市平均样本量更多，此时可以在一定程度上规避掉城市内部样本量过少的问题，但不可避免的妥协是城市数量的进一步减少。结果如表V3第11行所示，可以发现，即使在更少的样本量下，家政服务体系建设对女性的劳动参与和周工作时长的正面影响依然稳健。

**11. 反事实检验：家政服务体系建设对男性就业的影响**

在“男主外，女主内”的传统观念影响下，我国形成了母亲被视为“理想照顾者”、父亲被视为“理想工作者”的文化规则（陶艳兰和风笑天，2020）。家庭中主要是女性承担照料幼儿、老人与家务劳动，而男性主要负责养家糊口。因此，城市家政服务体系建设带来的家政服务业发展，减轻的主要是女性承担的照料活动与家务劳动负担，进而促进女性就业。现实中，我国男性较少从事家政服务工作。因此，预期城市家政服务体系建设应该不会影响当地男性的就业行为。这里使用男性样本，重新估计回归方程（2）式，如果发现男性就业行为也受到家政服务体系建设政策的影响，那么可能存在其他影响劳动力市场的混淆因素。表V4的估计结果显示，家政服务体系建设试点政策对男性的就业概率以及周工作小时数都没有显著的影响。

表V4 基于男性样本的反事实检验

|  |  |  |
| --- | --- | --- |
|  | (1) | (2) |
|  | 男性样本 | |
| 被解释变量 | 是否就业 | 周工作小时数 |
| 家政服务体系建设 | -0.0529 | -1.239 |
|  | (0.0403) | (2.199) |
| 样本量 | 3,568 | 3,427 |
| 调整R平方 | 0.387 | 0.318 |
| 个人与家庭特征 | 控制 | 控制 |
| 个体FE | 控制 | 控制 |
| 年份FE | 控制 | 控制 |
| 期初特征\*年份FE | 控制 | 控制 |

注: 括号中为bootstrap稳健标准误（抽样次数为1000），\*\*\*、\*\*、\*分别表示 1% 、5% 、10%的显著水平。个人与家庭特征包括：年龄及其平方项、受教育年限、是否在婚、是否农村户口、家庭成员数量。城市层面的期初特征包括：2009年人均GDP对数、2010年14岁以下少儿和65岁以上老年人占比、劳动年龄女性就业率。

**附录**VI 基于CHNS数据考察家政服务体系建设对高低学历女性的

异质影响

按照本文的论述，家政服务业发展从劳动力需求端促进低学历女性的工作参与，从服务供给端替代女性的家务活动和照料负担，从而促进她们的就业参与和工作时长。表VI1展示相应的考察结果。可以发现，家政服务体系建设使得初中及以下学历女性的劳动参与概率提高13.1%，每周工作时长增加5.3个小时（P值为0.106）。相比之下，家政服务体系建设使得高中及以上学历女性的劳动参与率提高9.96%，每周工作时长增加5.64个小时。

表VI1 对高低学历女性的异质性影响

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | 初中及以下学历女性 | | 高中及以上学历女性 | |
| 变量名称 | 劳动参与 | 周工作  时长(小时) | 劳动参与 | 周工作  时长(小时) |
| 家政服务体系建设 | 0.131\*\* | 5.373+ | 0.0996\* | 5.644\*\* |
|  | (0.0643) | (3.323) | (0.0520) | (2.569) |
| 样本量 | 1,936 | 1,875 | 1,788 | 1,753 |
| 调整R平方 | 0.704 | 0.687 | 0.811 | 0.782 |
| 个体FE | YES | YES | YES | YES |
| 年份FE | YES | YES | YES | YES |
| 选择变量\*年份FE | YES | YES | YES | YES |

注: 括号中为bootstrap稳健标准误（抽样次数为1000），\*\*\*、\*\*、\*分别表示 1% 、5% 、10%的显著水平。个人与家庭特征包括：年龄及其平方项、受教育年限、是否在婚、是否农村户口、家庭成员数量。城市层面的期初特征包括：2009年人均GDP对数、2010年14岁以下少儿和65岁以上老年人占比、劳动年龄女性就业率。

**附录**VII基于UHS数据考察家政服务体系建设对家政服务消费的

影响

这里尝试解决基于UHS数据考察时的交错型DID的估计偏误问题。由于UHS数据是混合截面数据，难以匹配个体的跨年度ID。也正是这个原因，使得当前主流的解决面板模型交错型DID的方法难以实施。笔者经过认真的研读相关文献和程序的help文档，发现这些方法的使用前提大都是——需要识别个体跨年度ID构成面板数据，甚至一些方法要求构成平衡面板数据。因此，我们无法在基于UHS数据的考察中采用当下流行的方法对估计偏误进行校正。为了应对交错型DID的潜在偏误，我们采用Cengiz（2019）提出的堆叠回归估计量（Stacked Regression Estimation）方法作为一种稳健性检验。正如刘冲等（2022）所述，目前对于堆叠回归统计软件只能估计动态效应的系数，难以给出平均处理效应，研究者多采用手动计算的方法。为避免手动堆叠带来的数据重复使用问题，本文具体操作如下：（1）筛选出2006-2014年UHS数据，该时间段UHS数据覆盖48个城市，进入处理组的有9个城市，其中2010年5个城市，2011年4个城市；（2）将2011年4个城市样本提出，对其年份和处理信息进行减1平移，由此获得与2010年对齐的处理组，将平移后的数据放入到原数据集（此时原数据集已经删除2011年4个城市数据），由此获得对齐的堆叠数据集；（3）基于对齐的数据进行双向固定效应回归，得到相关估计系数。结果如表VII1所示，可以发现解决交错型DID的潜在偏误后，结果与正文中表4结果保持一致。

表VII1 家政服务体系建设对家庭家政服务消费的影响-解决交错型did偏误

|  |  |  |
| --- | --- | --- |
|  | (1) | (2) |
|  | 妻子高中及以上学历 | 妻子初中及以下学历 |
|  | 基于混合截面数据的堆叠估计 | |
| 变量名称 | log家庭的家政服务消费 | |
| 家政服务体系建设 | 0.328\*\* | 0.0728 |
|  | (0.141) | (0.131) |
| 是否有70岁  以上老人 | 0.295\*\*\* | 0.0904 |
| (0.0738) | (0.0686) |
| 家政服务体系建设  \*有70岁以上老人 | 0.444\*\*\* | 0.172 |
| (0.161) | (0.188) |
| 是否有6岁  以下儿童 | 0.211\*\*\* | 0.0723\* |
| (0.0558) | (0.0410) |
| 家政服务体系建设  \*有6岁以下儿童 | -0.0261 | -0.0555 |
| (0.0909) | (0.178) |
| 样本量 | 31,973 | 16,455 |
| 调整R平方 | 0.218 | 0.190 |
| 个人与家庭特征 | 控制 | 控制 |
| 城市FE | 控制 | 控制 |
| 年份FE | 控制 | 控制 |
| 城市\*年份趋势 | 控制 | 控制 |
| 省份\*年份FE | 控制 | 控制 |
| 初期特征\*年份FE | 控制 | 控制 |

注: 括号中是聚类到城市层面的稳健标准误，\*\*\*、\*\*、\*分别表示 1% 、5% 、10%的显著水平。个人与家庭层面变量包括：户主是否男性、丈夫年龄、丈夫受教育年数、妻子年龄、妻子受教育年数、家庭成员数量、家庭收入的对数。城市层面的期初特征包括：2009年人均GDP对数、2010年14岁以下少儿和65岁以上老年人口占比、劳动年龄女性就业率。

参考文献

1. Borusyak K., X. Jaravel, and J. Spiess, “Revisiting Event Study Designs: Robust and Efficient Estimation”, 2021, Working Papers.
2. Callaway, B., and PHC. Sant’ Anna, “Difference-in-Differences with Multiple Time Periods” ,*Journal of Econometrics*, 2021, 225(2), 200-230.
3. Cengiz, D., A. Dube, A. Lindner, and B. Zipperer, “The Effect of Minimum Wages on Low-Wage Jobs”, *The Quarterly Journal of Economics*, 2019,134(3), 1405-1454.
4. Chetty, R., A. Looney, and K. Kroft, "Salience and taxation: Theory and evidence."*American Economic Review*, 2009, 99 (4), 1145-77.
5. Chyi, H., and B. Zhou, “The Effects of Tuition Reforms on School Enrollment in Rural China”, *Economics of Education Review*, 2014, 38, 104-123.
6. De Chaisemartin, C., X. D’Haultfoeuille, “Two-way Fixed Effects Regressions with Several Treatments”, Working Paper, 2020.
7. 范红丽、辛宝英，“家庭老年照料与农村妇女非农就业——来自中国微观调查数据的经验分析”，《中国农村经济》，2019 年第2期，第98-114页。
8. Fang, G., and H. Cao, “State versus Private Provision: How Does China’s Market‐Oriented Reform Affect Healthcare Delivery?”, *Economics of Transition and Institutional Change*, 2020, 28(3), 381-411.
9. Gardner J., “Two-Stage Differences in Differences”, 2021,Working Paper.
10. Goodman-Bacon, A. “Difference-in-differences with Variation in Treatment Timing”, *Journal of Econometrics*, 2021, 225(2), 254-277.
11. Li, P., Y. Lu, and J. Wang, “Does Flattening Government Improve Economic Performance? Evidence from China” , *Journal of Development Economics*, 2016, 123, 18-37.
12. 刘冲、沙学康、张妍，“交错双重差分：处理效应异质性与估计方法选择”，《数量经济技术经济研究》，2022年第9期，第177-204页。
13. Sun L., and S. Abraham, “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects”, *Journal of Econometrics*, 2020, 225(2), 175-199.
14. 陶艳兰、风笑天， “‘理想照顾者’文化规则的破解：社会政策支持母亲就业的关键问题”，《社会科学》，2020年第4期，第77-89页。
15. 熊瑞祥、李辉文，“儿童照管、公共服务与农村已婚女性非农就业——来自CFPS数据的证据”，《经济学》(季刊)，2016年第1期，第393-414页。
16. 续继、黄娅娜，“性别认同与家庭中的婚姻及劳动表现”，《经济研究》，2018年第4期，第136-150页。
17. 张川川、王靖雯，“性别角色与女性劳动力市场表现”，《经济学》(季刊)，2020年第3期，第977-994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. 感谢审稿专家的建设性修改意见。 [↑](#footnote-ref-0)
2. 在基本回归模型中未控制家庭人均收入的对数，是因为女性的就业参与行为本身会影响家庭收入，因此在基本回归模型中控制该变量，会带来“坏的控制变量”（badcontrol）问题（Angrist and Pischke, 2009）。 [↑](#footnote-ref-1)
3. 由于缺乏城市层面历年养老院床位数量的官方统计，故我们在“中国养老网”上获取了2020年我国各城市的养老院床位数，由此得到各城市每万常住人口养老院床位数。 [↑](#footnote-ref-2)
4. 感谢审稿专家的建设性意见。 [↑](#footnote-ref-3)