**公共卫生事件对流动儿童和本地儿童非认知能力的影响**

卓 依 张 竞 冯帅章 陈媛媛

**目录**

[附录I 疫情对学习表现和身体健康等其他人力资本的影响 1](#_Toc170660295)

[附录II 附图 3](#_Toc170660296)

[附录III 稳健性检验 5](#_Toc170660297)

[附录IV 样本选择偏误问题 8](#_Toc170660298)

[附录V 机制分析：社会环境 10](#_Toc170660299)

[附录VI 疫情背景下非认知能力差异的分解 11](#_Toc170660300)

[参考文献 13](#_Toc170660301)

# 附录I 疫情对学习表现和身体健康等其他人力资本的影响

为更全面地展示疫情对于学生整体人力资本发展的影响，此附录将提供关于疫情对两类儿童的学习表现和身体健康等其他人力资本影响差异的简单讨论。

首先，我们使用2020年调查数据中班主任对每个学生在疫情期间的出勤情况、作业完成情况、上课遵守纪律情况和学习成绩等方面的评价来检验两类儿童在疫情期间的学习表现差异。[[1]](#footnote-1)表I1第（1）-（4）列汇报了回归结果，总体上并未发现两类儿童在学习表现上存在显著差异。[[2]](#footnote-2)尽管使用老师的评价存在一定的主观性，但以上的分析至少可以说明疫情并没有显著改变两类儿童学业表现的差距。

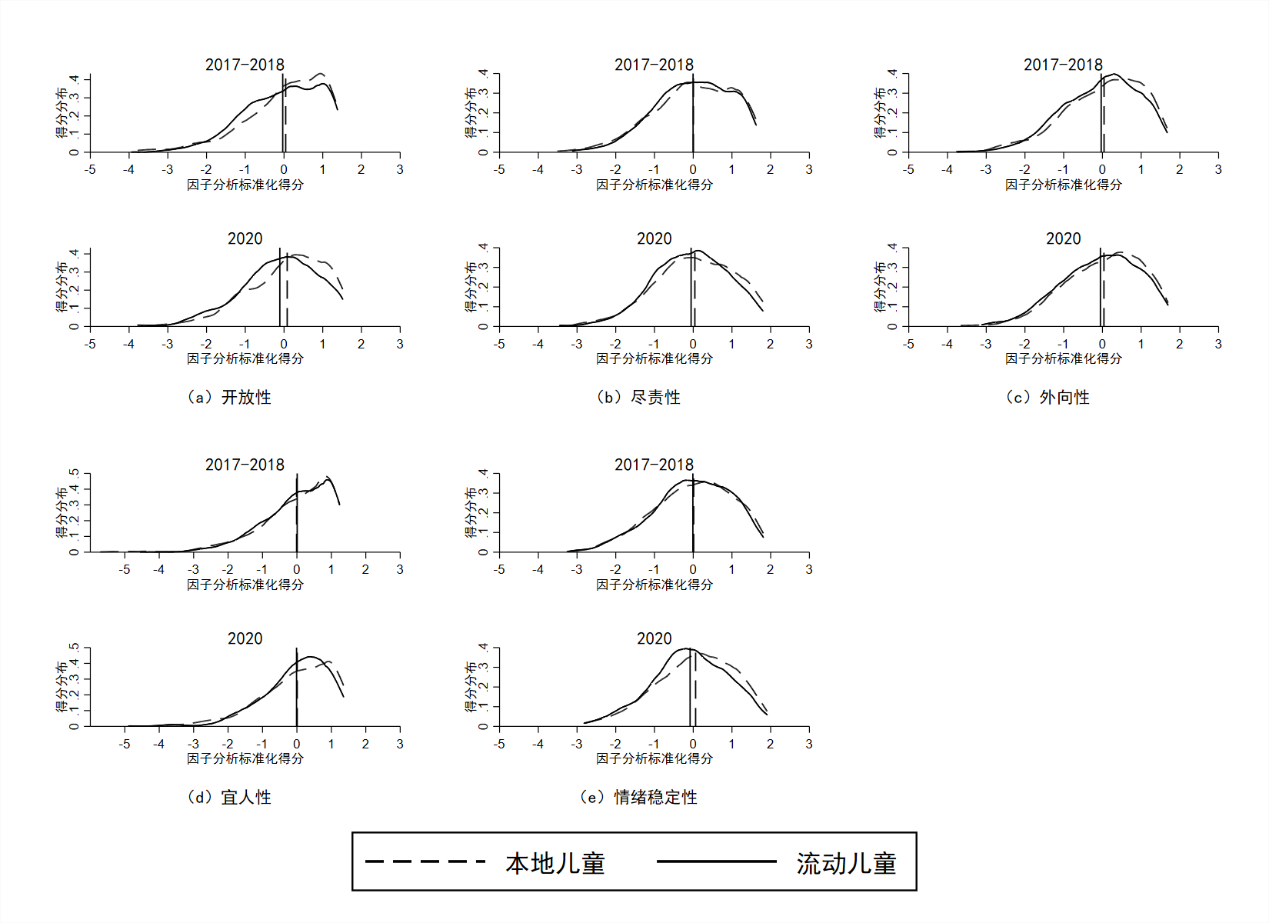
其次，疫情可能会对儿童的身体健康产生影响。在网课期间，由于不能外出，锻炼时间较少，长时间接触电子设备也可能会使学生视力下降。表I1第（5）-（7）列分别对学生自评和家长评价的健康程度，以及是否有近视加深的情况进行分析，并未发现两类儿童在身体健康和视力变化等方面存在显著差异。尽管现有文献发现身体健康会对非认知能力产生影响（Elkins et al., 2017），但以上结果至少表明两类儿童在情绪稳定性上的差异并不是健康状况造成的。

表I1 疫情期间本地儿童和流动儿童在其他人力资本上的差异

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **变量** | **学习表现** | | | | **健康状况** | | |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** |
| **出勤情况** | **作业完成情况** | **上课遵守纪律情况** | **平时成绩** | **学生自评健康** | **家长评价的学生健康** | **近视加深** |
| 流动儿童 | 0.004 | -0.011 | 0.019 | -0.015 | -0.017 | -0.006 | 0.008 |
|  | (0.031) | (0.022) | (0.019) | (0.048) | (0.017) | (0.010) | (0.030) |
| 出勤情况 | 0.904\*\*\* |  |  |  |  |  |  |
| （上学期） | (0.059) |  |  |  |  |  |  |
| 作业完成情况 |  | 0.869\*\*\* |  |  |  |  |  |
| （上学期） |  | (0.036) |  |  |  |  |  |
| 上课遵守纪律情况 |  |  | 0.863\*\*\* |  |  |  |  |
| （上学期） |  |  | (0.037) |  |  |  |  |
| 学校-年级固定效应 | 是 | 是 | 是 | 是 | 是 | 是 | 是 |
| 控制变量 | 是 | 是 | 是 | 是 | 是 | 是 | 是 |
| 观测值 | 1,288 | 1,288 | 1,288 | 1,288 | 1,288 | 1,288 | 1,288 |
| R-squared | 0.401 | 0.531 | 0.688 | 0.113 | 0.017 | 0.020 | 0.031 |

注： “出勤情况”“作业完成情况”“上课遵守纪律情况”和“平时成绩”，由班主任分“一般”“良好”和“优异”三个等级进行评价。“自评健康良好”和“家长评价健康良好”分别指学生自评和家长评价的学生健康良好的虚拟变量，即疫情期间没有生病；“近视加深”指家长报告的疫情期间学生视力下降的虚拟变量。控制变量包括性别、独生子女、父亲/母亲受教育水平在高中及以上、家庭月平均收入在1万元及以上。标准误聚类在学校-年级-班级层面。\*\*\*、\*\*、\*分别表示在1%、5%、10%水平上统计显著。

# 附录II 附图



图A1 疫情前后本地儿童和流动儿童的大五人格得分分布

图A2 本地儿童和流动儿童家长的职业分布（分家长教育水平）

注：左轴纵坐标代表相应职业在2020年出现工资下降的比例，即表示对应职业受疫情影响的严重程度。右轴纵坐标代表本地儿童或流动儿童家长在对应职业上的人数占比。

# 附录III 稳健性检验

表III1 展示了一系列稳健性检验结果。首先，变换流动儿童的定义，重新对基准回归进行估计，结果展示在Panel A和Panel B中。在基准回归中流动儿童的定义为“不具有上海户籍的儿童”，这与通常人们所指的低技能、低收入家庭的流动儿童并不等同。Panel A 将流动儿童定义为“具有农村户口的非上海户籍儿童”[[3]](#footnote-3)，即非上海户籍儿童样本仅保留具有农村户口的儿童样本，占非户籍儿童总数的51%；Panel B考虑到目前城市的人口已经从简单的二元结构（即本地居民和流动人口）扩展为本地人口、外来高技能劳动力和外来普通劳动力的新三元结构，且后两者之间存在较为显著的差异（梁文泉和陆铭，2015），因此将流动儿童定义为“父亲和母亲均没有大学以上学历的非上海户籍儿童”，占非上海户籍儿童总数的86%。在新的定义之下，Panel A和Panel B的估计样本中均删除了其他非上海户籍儿童样本。总的来看，不同流动儿童定义下的结果基本稳健**。**

第二，将民办农民工子弟学校的流动儿童样本纳入，从广泛的角度讨论疫情对流动儿童的影响。在基准回归中，为排除学生进入不同学校类型的选择性偏差，只使用公办学校的数据，但考虑到通常来自民办农民工子弟学校的流动儿童表现会相比公办学校的更差一些，只使用公办学校的数据，可能会低估疫情对流动儿童的负面影响。为此，Panel C同时使用公办学校和民办学校的学生样本。考虑到民办农民工子弟学校中只包括流动儿童，因此无法使用学校固定效应进行估计。为此，模型中加入公办学校的虚拟变量，及其与控制变量的交互项，重新估计疫情对两类儿童的影响。结果显示流动儿童与疫情后交互项的系数均为负，且在开放性和情绪稳定性维度上显著，说明同时考虑了两类学校的流动儿童的结果依然稳健。

第三，采用不同方式来构建大五人格。Panel D是对44个问题先按年份进行标准化后，再进行因子分析。Panel E、F和G则是在使用因子分析生成大五人格后，采用不同方式对其进行标准化，分别为将三年数据混合进行标准化，分各年份进行标准化，以及按照性别-年级进行标准化。[[4]](#footnote-4)以上采用不同方式来生成或标准化大五人格的结果均与基准回归一致。

第四，调整基准方程或分析样本，重新进行估计。Panel H考虑学生的非认知能力发展与其年龄、家长的年龄密切相关，在基准方程中加入学生的月龄和月龄的平方、家长的年龄和年龄的平方[[5]](#footnote-5)；Panel I考虑到2018年数据是2017年数据的追踪样本，为使疫情前后的数据更具可比性，去除2018年的数据重新估计基准回归方程。估计结果均与基准回归结果保持一致，且开放性的交互项系数在10%的水平下显著。

表III1 稳健性检验

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **变量** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
| **开放性** | **尽责性** | **外向性** | **宜人性** | **情绪稳定性** |
| **Panel A 流动儿童定义：农村户口的非户籍儿童** | | | | | |
| 流动儿童×疫情后 | -0.151 | -0.147 | -0.224 | -0.200 | -0.416\*\*\* |
|  | (0.140) | (0.121) | (0.178) | (0.197) | (0.147) |
| 观测值 | 2,622 | 2,622 | 2,622 | 2,622 | 2,622 |
| R-squared | 0.104 | 0.108 | 0.074 | 0.092 | 0.076 |
| **Panel B 流动儿童定义：父亲和母亲均没有大学以上学历的非户籍儿童** | | | | | |
| 流动儿童×疫情后 | -0.186\* | -0.174\* | -0.138 | -0.127 | -0.313\*\*\* |
|  | (0.106) | (0.102) | (0.105) | (0.102) | (0.103) |
| 观测值 | 3,252 | 3,252 | 3,252 | 3,252 | 3,252 |
| R-squared | 0.099 | 0.104 | 0.068 | 0.082 | 0.073 |
| **Panel C 加入民办农民工子弟小学的样本** | | | | | |
| 流动儿童×疫情后 | -0.180\* | -0.142 | -0.131 | -0.097 | -0.185\*\* |
|  | (0.096) | (0.097) | (0.093) | (0.086) | (0.090) |
| 观测值 | 7,069 | 7,069 | 7,069 | 7,069 | 7,069 |
| R-squared | 0.038 | 0.024 | 0.035 | 0.038 | 0.029 |
| **Panel D 大五构建：先标准化每个问题再进行因子分析** | | | | | |
| 流动儿童×疫情后 | -0.174 | -0.139 | -0.164 | -0.142 | -0.260\*\* |
|  | (0.106) | (0.101) | (0.118) | (0.104) | (0.108) |
| 观测值 | 3,510 | 3,510 | 3,510 | 3,510 | 3,510 |
| R-squared | 0.097 | 0.101 | 0.070 | 0.079 | 0.071 |
| **Panel E 大五构建：三年混合标准化** | | | | | |
| 流动儿童×疫情后 | -0.162 | -0.127 | -0.140 | -0.123 | -0.222\*\* |
|  | (0.098) | (0.092) | (0.100) | (0.090) | (0.092) |
| 观测值 | 3,510 | 3,510 | 3,510 | 3,510 | 3,510 |
| R-squared | 0.097 | 0.101 | 0.070 | 0.079 | 0.071 |
| **Panel F 大五构建：分年份进行标准化** | | | | | |
| 流动儿童×疫情后 | -0.162 | -0.127 | -0.137 | -0.124 | -0.222\*\* |
|  | (0.098) | (0.092) | (0.101) | (0.090) | (0.093) |
| 观测值 | 3,510 | 3,510 | 3,510 | 3,510 | 3,510 |
| R-squared | 0.097 | 0.100 | 0.070 | 0.079 | 0.071 |

（续下表）

表III1 稳健性检验（续表）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **变量** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
| **开放性** | **尽责性** | **外向性** | **宜人性** | **情绪稳定性** |
| **Panel G 大五构建：分性别-年级进行标准化** | | | | | |
| 流动儿童×疫情后 | -0.162 | -0.129 | -0.137 | -0.121 | -0.220\*\* |
|  | (0.099) | (0.094) | (0.100) | (0.090) | (0.092) |
| 观测值 | 3,510 | 3,510 | 3,510 | 3,510 | 3,510 |
| R-squared | 0.087 | 0.091 | 0.068 | 0.067 | 0.068 |
| **Panel H 控制学生和家长年龄** | | |  |  |  |
| 流动儿童×疫情后 | -0.201\* | -0.132 | -0.115 | -0.095 | -0.239\*\* |
|  | (0.106) | (0.097) | (0.113) | (0.107) | (0.113) |
| 观测值 | 2,727 | 2,727 | 2,727 | 2,727 | 2,727 |
| R-squared | 0.118 | 0.124 | 0.083 | 0.096 | 0.086 |
| **Panel I 去掉2018年样本** |  |  |  |  |  |
| 流动儿童×疫情后 | -0.194\* | -0.157 | -0.165 | -0.149 | -0.198\* |
|  | (0.109) | (0.104) | (0.111) | (0.098) | (0.105) |
| 观测值 | 2,788 | 2,788 | 2,788 | 2,788 | 2,788 |
| R-squared | 0.093 | 0.104 | 0.068 | 0.079 | 0.070 |
| 控制变量 | 是 | 是 | 是 | 是 | 是 |
| 控制变量×疫情后 | 是 | 是 | 是 | 是 | 是 |
| 控制变量×流动儿童 | 是 | 是 | 是 | 是 | 是 |

注：Panel C还控制了流动儿童、公办学校、控制变量×公办学校交互项、年份和年级固定效应。其他Panel的模型与基准回归一致，控制了相关的控制变量、控制变量×疫情后交互项、控制变量×流动儿童交互项，以及年份-学校-年级固定效应。其中，控制变量包括性别、独生子女、父亲/母亲受教育水平在高中及以上、家庭月平均收入在1万元及以上。所有回归的标准误都聚类在年份-学校-年级-班级层面。\*\*\*、\*\*、\*分别表示在1%、5%、10%水平上统计显著。

# 附录IV 样本选择偏误问题

考虑到疫情可能导致部分流动人口因无法复工复产而选择不回上海，2020年调查样本学校中可能存在部分流动儿童离开上海的情况，导致调查样本的非随机缺失，进而产生样本自选择的偏误问题（selection bias）。2020年学校问卷的访问结果显示，样本学校学生流失率在2.4%-4.8%之间，其中四、五年级的流失率分别为4%和3%。2017、2018和2020年样本学校四、五年级的平均流失率分别是8.5%、3.7%和3.5%，2020年样本学校的平均流失率相比2017和2018年并没有上升的趋势，因此，2020年流动儿童的流失很可能主要不是受到疫情的影响，而是与疫情前一样，主要由异地中考政策导致的。

为进一步检验样本选择可能导致的偏误，本文参考Lee（2009）的方法，计算不同的偏误方向对结果产生的可能影响。具体而言，如果是那些家庭背景更差、非认知能力更低的流动儿童在疫情后更可能回到老家，则基准回归结果将低估疫情对两类儿童非认知能力差距的影响；反之，如果疫情导致家庭背景更好、非认知能力更高的流动儿童在疫情后更可能回老家，则估计结果将高估疫情的影响。从直觉上看，疫情通常对低技能劳动力的就业冲击更大，因此本文的基准结果更有可能因为样本的选择问题而被低估。

参考Lee（2009）的方法，假设在极端情况下，那些家庭背景最差、非认知能力最低的流动儿童最可能受疫情影响回老家，然后分别按样本学校学生流失率2.4%、4.8%，以及按四年级4%和五年级3%比例，复制对应比例的位于非认知能力分布尾端的样本，并补齐至原样本中，重新估计基准回归，得到疫情对两类儿童情绪稳定性差距影响的上界（绝对值），即疫情对两类儿童非认知能力差距的最大影响。[[6]](#footnote-6)如表IV1第（3）（5）和（7）列所示，在不同流失率的假设下，修正样本后的交互项系数分别为-0.276、-0.310和-0.294个标准差，且均在1%水平上显著。反之，假设那些家庭背景最好、非认知能力最高的流动儿童最可能回老家，也按相应的流失率复制对应比例的位于非认知能力分布前端的样本，重新估计基准回归，得到疫情影响的下界（绝对值）。如表IV1第（2）（4）和（6）列所示，在不同的流失率下，系数估计的下界分别为-0.187、-0.143和-0.161个标准差，其中在2.4%和分年级（四年级4%，五年级3%）比例时，系数仍在10%水平下显著。

表IV1 参考Lee（2009）计算基准结果上界和下界

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **变量** | **情绪稳定性** | | | | | | |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** |
| **基准回归** | **下界\_2.4%** | **上界\_2.4%** | **下界\_4.8%** | **上界\_4.8%** | **下界\_分年级** | **上界\_分年级** |
| 流动儿童×疫情后 | -0.222\*\* | -0.187\* | -0.276\*\*\* | -0.143 | -0.310\*\*\* | -0.161\* | -0.294\*\*\* |
|  | (0.093) | (0.095) | (0.092) | (0.096) | (0.096) | (0.096) | (0.093) |
| 流动儿童 | 0.129 | 0.125 | 0.118 | 0.114 | 0.123 | 0.121 | 0.123 |
|  | (0.106) | (0.107) | (0.107) | (0.107) | (0.107) | (0.107) | (0.107) |
| 年份-学校-年级固定效应 | 是 | 是 | 是 | 是 | 是 | 是 | 是 |
| 控制变量 | 是 | 是 | 是 | 是 | 是 | 是 | 是 |
| 控制变量×疫情后 | 是 | 是 | 是 | 是 | 是 | 是 | 是 |
| 控制变量×流动儿童 | 是 | 是 | 是 | 是 | 是 | 是 | 是 |
| 观测值 | 3,510 | 3,524 | 3,523 | 3,537 | 3,536 | 3,531 | 3,529 |
| R-squared | 0.071 | 0.072 | 0.072 | 0.073 | 0.074 | 0.072 | 0.072 |

注：第（6）和（7）列的“分年级”指按照四年级流失率4%和五年级3%的比例分别进行计算。控制变量包括性别、独生子女、父亲/母亲受教育水平在高中及以上、家庭月平均收入在1万元及以上。标准误聚类在年份-学校-年级-班级层面。\*\*\*、\*\*、\*分别表示在1%、5%、10%水平上统计显著。

# 附录V 机制分析：社会环境

为讨论来自疫情越严重地区的流动儿童是否受到疫情的影响越大，本文使用2020年的流动儿童样本，运用以下方程进行估计：

其中，代表不同计算节点下来源地省份*p*累计确诊人数的log值。具体采取以下4个时间节点：从疫情爆发至调查时间（即疫情爆发至6月15日）、网课开始至调查时间（3月2日至6月15日）、调查前一个半月（5月1日至6月15日）和调查前一个月（5月15日至6月15日）。考虑到疫情爆发初期湖北（主要是武汉地区）的严重疫情可能会给来自湖北的流动儿童（占样本的4%）带来较大的心理冲击，我们进一步在方程中加入湖北省的虚拟变量（即）。

表V1 结果显示，在不同的计算节点下，户籍所在地的疫情相对严重程度对流动儿童情绪稳定性的影响均很小且不显著。同时，前面的系数也不显著，且均为正。总的来说，我们并未发现来源地疫情严重程度会对流动儿童的情绪稳定性造成显著负向影响。

表V1 户籍所在地疫情严重程度对流动儿童情绪稳定性的影响

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **变量** | **情绪稳定性** | | | |
| **(1)** | **(2)** | **(3)** | **(4)** |
| *Log\_cases*（疫情爆发-调查时间） | -0.062 |  |  |  |
|  | (0.091) |  |  |  |
| *Log\_cases*（网课开始-调查时间） |  | 0.019 |  |  |
|  |  | (0.029) |  |  |
| *Log\_cases*（调查前一个半月） |  |  | 0.049 |  |
|  |  |  | (0.043) |  |
| *Log\_cases*（调查前一个月） |  |  |  | 0.055 |
|  |  |  |  | (0.043) |
| 是否来自湖北省（1=是） | 0.470 | 0.107 | 0.121 | 0.186 |
|  | (0.485) | (0.281) | (0.247) | (0.235) |
| 学校-年级固定效应 | 是 | 是 | 是 | 是 |
| 控制变量 | 是 | 是 | 是 | 是 |
| 观测值 | 559 | 559 | 559 | 559 |
| R-squared | 0.053 | 0.053 | 0.055 | 0.055 |

注：控制变量包括性别、独生子女、父亲/母亲受教育水平在高中及以上、家庭月平均收入在1万元及以上。标准误聚类在学校-年级-班级层面。\*\*\*、\*\*、\*分别表示在1%、5%、10%水平上统计显著。

# 附录VI 疫情背景下非认知能力差异的分解

机制分析发现疫情期间本地儿童和流动儿童的情绪稳定性差异主要由家庭经济生活情况与家庭学习环境两个方面因素导致的，为比较不同因素影响的差异，本文通过分解的方法进行分析。首先，修改基准方程（1）为方程（Ⅵ1），使用2020年数据检验疫情期间流动儿童和本地儿童的情绪稳定性差异：

（Ⅵ1）。

再依次将方程（3）的机制变量放入到方程（Ⅵ2）：

（Ⅵ2）。

表VI1第（1）列在不控制任何机制变量，仅控制基准回归中与家庭背景相关的特征变量和学校-年级固定效应的情况下，流动儿童和本地儿童的情绪稳定性差距为0.116个标准差，且在10%水平上显著。第（2）-（5）列分别加入各类机制变量，结果显示，流动儿童前面的系数均变小且不显著，家庭经济困难与学生的情绪稳定性存在显著负相关关系，而学习环境安静、家长查看老师对孩子作业的批改和反馈两个因素与学生的情绪稳定性存在显著正相关关系。第（6）列将所有的机制变量放入方程中，流动儿童的系数降为-0.038个标准差。

结合方程（3）、（Ⅵ1）和（Ⅵ2），并参考Gelbach（2016）的分解方法[[7]](#footnote-7)，可将两类儿童的情绪稳定性差异表示如下：

*=+*  ，

其中，代表流动儿童和本地儿童的情绪稳定性差异，其由两部分构成：一部分是两类儿童无法被观测到的差异，即表VI1第（6）列流动儿童变量的系数；另一部分是可以由各个机制变量所解释的部分，每个机制变量可解释的部分为，其中为方程（Ⅵ2）中前面的系数，估计结果显示在表VI1第（6）列；而为方程（3）中流动儿童变量前面的估计系数，各机制变量对应的估计系数显示在正文表4、表5和表6中。基于此，每个机制变量对于两类儿童的情绪稳定性差异的解释力度可计算为 。

表VI1 机制分析回归结果

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **被解释变量** | **情绪稳定性** | | | | | |
| **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| 流动儿童 | -0.116\* | -0.088 | -0.100 | -0.081 | -0.105 | -0.038 |
|  | (0.065) | (0.070) | (0.064) | (0.062) | (0.066) | (0.068) |
| 收入下降 |  | -0.062 |  |  |  | -0.050 |
|  |  | (0.055) |  |  |  | (0.053) |
| 经济困难 |  | -0.143\*\* |  |  |  | -0.091 |
|  |  | (0.068) |  |  |  | (0.068) |
| 父亲焦虑 |  |  | -0.121 |  |  | -0.099 |
|  |  |  | (0.074) |  |  | (0.075) |
| 母亲焦虑 |  |  | -0.046 |  |  | -0.010 |
|  |  |  | (0.062) |  |  | (0.065) |
| 学习环境安静 |  |  |  | 0.497\*\*\* |  | 0.485\*\*\* |
|  |  |  |  | (0.073) |  | (0.073) |
| 下载和上传作业 |  |  |  |  | -0.017 | -0.004 |
|  |  |  |  |  | (0.063) | (0.065) |
| 辅导作业和解答问题 |  |  |  |  | 0.033 | 0.039 |
|  |  |  |  |  | (0.059) | (0.058) |
| 查看老师对孩子作业的 |  |  |  |  | 0.117\*\* | 0.114\*\* |
| 批改和反馈 |  |  |  |  | (0.057) | (0.055) |
| 学校-年级固定效应 | 是 | 是 | 是 | 是 | 是 | 是 |
| 控制变量 | 是 | 是 | 是 | 是 | 是 | 是 |
| 观测值 | 1,288 | 1,288 | 1,288 | 1,288 | 1,288 | 1,288 |
| R-squared | 0.057 | 0.062 | 0.063 | 0.086 | 0.061 | 0.096 |

注：控制变量包括性别、独生子女、父亲/母亲受教育水平在高中及以上、家庭月平均收入在1万元及以上。标准误聚类在学校-年级-班级层面。\*\*\*、\*\*、\*分别表示在1%、5%、10%水平上统计显著。

# 参考文献

1. Cook, C., R. Diamond, J. V. Hall, J. A. List, and P. Oyer, “The Gender Earnings Gap in the Gig Economy: Evidence from Over a Million Rideshare Drivers”, *The Review of Economic Studies*, 2021, 88(5), 2210-2238.
2. Elkins, R. K., S. C. Kassenboehmer, and S. Schurer, “The Stability of Personality Traits in Adolescence and Young Adulthood”, *Journal of Economic Psychology*, 2017, 60, 37-52.
3. Gelbach, J. B., “When Do Covariates Matter? And Which Ones, and How Much?”, *Journal of Labor Economics*, 2016, 34 (2), 509-543.
4. Lee, D. S., “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects”, *Review of Economic Studies*, 2009, 76 (3), 1071-1102.
5. 梁文泉、陆铭，“城市人力资本的分化：探索不同技能劳动者的互补和空间集聚”，《经济社会体制比较》，2015年第3期，第185-197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. 最优的情况是可以使用标准化的期末考试成绩来衡量学生的学习表现，但由于考虑到网课的上课质量以及学生的心理承受能力，上海市教育局全面取消了小学2020年春季学期期末考试。 [↑](#footnote-ref-1)
2. 考虑到班主任对以上几个方面的评价为有序离散变量，本文也使用Ordered Probit模型进行估计，结果基本一致。 [↑](#footnote-ref-2)
3. 2020年个别学校五年级学生没有城市或农村户口信息，因此使用该信息会导致回归样本有一定减少。 [↑](#footnote-ref-3)
4. 由于 2020 年调查学生的出生年月有部分学校信息缺失，又考虑到学生均来自四、五年级，同一年级学生的年龄相近，因此按照性别-年级进行标准化。 [↑](#footnote-ref-4)
5. 需要说明的是，家长年龄和学生出生年月信息有部分缺失，其中2017-2018年这两个变量的缺失样本占基准回归分析样本的18%，2020年的缺失样本占28%。因此，这里用于回归分析的样本有所减少。 [↑](#footnote-ref-5)
6. 这里假设疫情的影响是单调的，即疫情的影响只能要么是使得流动儿童更多回到老家，要么则是使得更多流动儿童回到上海。 [↑](#footnote-ref-6)
7. 文献中该方法主要运用于解释性别工资差距的问题（Cook et al., 2021；Gelbach, 2016）。 [↑](#footnote-ref-7)