

附录一：Pearson 检验方法

对于任意学校 $s=1, \dots, S$, 检验公式为:

$$T = \sum_c \sum_j (n_{cj} - \hat{n}_{cj})^2 / \hat{n}_{cj} \quad n_{c.} = \sum_j n_{cj}, \quad n_{.j} = \sum_c n_{cj}, \quad \hat{n}_{cj} = n_{c.} n_{.j} / \sum_c \sum_j n_{cj} \quad (1)$$

其中, n_{cj} 是学校 s 内部班级 $c = 1, \dots, C$ 内具有特征 $j = 1, \dots, J$ 的学生个数; \hat{n}_{cj} 是在特征因素和班级相互独立情况下, 预测所得的班级 c 内具有特征 j 的学生个数。在相互独立原假设条件下, T 服从自由度为 $(C-1)(J-1)$ 的卡方分布。进一步, 假设这 S 所学校是相互独立的, 将 S 个 T 统计值简单相加可得到一个自由度为 $[\sum_s (C-1)](J-1)$ 的总和统计值。

附录二：Ohinata 和 Van Ours(2013)检验方法

计算学校内部班级间留守儿童相差 d 个的概率:

假设学校内部留守儿童的总人数 n , 当 n 为偶数时

$$p(d=0) = \binom{n}{n/2} (1/2)^n, \quad (2)$$

$$p(d=2k | k \in N^*, k, n/2) = 2 \binom{n}{n/2-k} (1/2)^n \quad (d..2) \quad (3)$$

当 n 为奇数时, 对 $\forall d$

$$p(d=2k+1 | k \in N^*, k, (n-1)/2) = 2 \binom{n}{(n-1)/2-k} (1/2)^n \quad (4)$$

根据上述公式计算所得的概率和拥有不同数量留守儿童的学校个数, 可计算出随机分配下学校内部班级间留守儿童相差 d 个的学校数量。

附录三：班级构成随机性检验

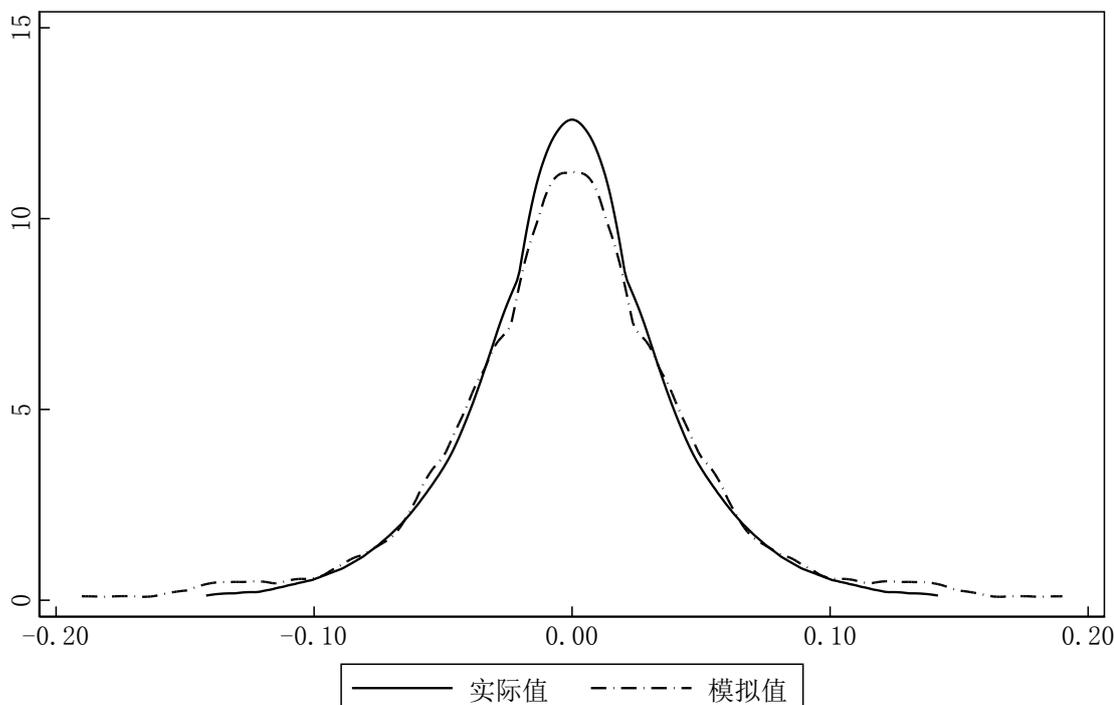
第一, 参考 Guryan 等(2009)的方法, 考察在年级 \times 学校固定效应下, 学生 i 的性别、年龄、留守状态、家庭经济状况和六年级自评成绩与同班同学的相应平均特征是否具有显著的相关性。由于班级的构成是在年级内部分配的结果, 因此本文在检验中进一步控制了学生 i 同年级同学的相应平均特征, 以此来纠正学生自身的特征与同班同学平均特征之间的机械负相关偏差。学生自评六年级自评成绩为语数英三科学习成绩的均值, 各科成绩均采用四分类李克特量表法: 特别吃力=1, ..., 一点也不吃力=4。附表 1 回归结果显示, 纠正偏差后学生 i 的留守状态等五种特征与同班同学平均特征之间不存在显著的相关性, 表明学校内部班级构成是随机的。

附表1 班级构成随机性检验：OLS模型

	性别	年龄	留守状态	家庭经济状况	六年级自评成绩
同班同学特征	-0.1137 (-0.9673)	0.0728 (1.3272)	0.0449 (1.3035)	0.0074 (0.1498)	0.0705 (1.6039)
同年级同学特征	-52.4198*** (-15.6766)	-50.7248*** (-14.5489)	-52.8585*** (-12.3751)	-53.6064*** (-20.1036)	-49.5290*** (-14.2607)
年级×学校固定效应	是	是	是	是	是
样本量	3089	3017	3089	3072	2995
调整的R ²	0.8619	0.9650	0.9238	0.8994	0.8377

注：*、**、***分别表示在10%、5%、1%的水平下显著。括号内为t统计值。下同。

第二，参照 Bietenbeck (2019)，采用 Monte Carlo 模拟法对留守儿童在学校内部班级间的分配是否具有随机性进行了进一步的检验。具体方法为模拟 1000 次在随机分配条件下，班级留守儿童比重对学校×年级固定效应回归后的残差分布，其中班级规模和年级内留守儿童的数量为回归样本的实际值，并将模拟的残差分布与实际残差分布进行对比，若不存在显著差异则证实留守儿童在班级间的分配是随机的。附图 1 为残差模拟分布与实际分布状况。从图中可以看出残差的模拟值和实际值的分布基本重合，Epps-Singleton 检验结果为 $W_2=0.274$ ， $p=0.9914$ ，表明两个分布之间不存在显著的差异，留守儿童在学校内部班级间的分配是随机的。



附图1 学校×年级固定效应下留守儿童比重的实际分布和模拟分布状况

附表2 班级构成 Pearson 卡方检验

	七年级	八年级	九年级
性别	0.9964	0.8085	0.9059
年龄	0.5020	0.3733	0.1807
留守儿童	0.3976	0.5790	0.4523
认知能力	0.1834	0.1109	0.3306
家庭经济状况	0.0766	0.7477	0.1490

附表3 班级间教师资源分配差异显著性检验：Mann-Whitney 检验

	班主任教师		语文教师		数学教师		英语教师	
	z值	p值	z值	p值	z值	p值	z值	p值
性别（男=1）	-0.4325	0.6654	-0.9043	0.3658	-0.8072	0.4195	0.8346	0.4039
受教育年限	-0.0480	0.9617	0.8368	0.4027	0.6936	0.4879	0.9127	0.3614
工作年限	0.3832	0.7016	0.4182	0.6758	0.1768	0.8597	-1.1992	0.2304

附表4 留守儿童和非留守儿童学习成绩

	留守儿童		非留守儿童	
	均值	标准差	均值	标准差
语文	69.8843	9.8332	70.0655	9.8502
数学	70.0487	9.8615	70.1381	9.9013
英语	70.0218	9.5480	70.0237	9.9239

附表5 控制变量定义和描述性统计

变量	定义	均值/百分比	标准差
个体特征			
留守儿童	留守儿童=1, 非留守儿童=0	11.65%	
性别	男=1, 女=0	51.54%	
年龄	年	14.01	1.36

农业户口	是=1, 否=0	63.54%	
本地户口	是=1, 否=0	73.03%	
健康	健康状况, 5分类: 非常不健康=1, ..., 非常健康=5	4.09	0.91
认知能力	认知能力水平测试标准化得分	-0.24	0.83
信心	对未来的信心程度, 4分类: 非常没信心=1, ..., 非常有信心=4	3.16	0.74
自评语文成绩	6年级语文学习状况, 4分类: 特别吃力=1, ..., 一点也不吃力=4	3.06	0.84
自评数学成绩	6年级数学学习状况, 4分类: 特别吃力=1, ..., 一点也不吃力=4	2.76	1.02
自评英语成绩	6年级英语学习状况, 4分类: 特别吃力=1, ..., 一点也不吃力=4	2.62	1.11
语文看法	语文对未来很有帮助, 4分类: 完全不同意=1, ..., 完全同意=4	3.34	0.78
数学看法	数学对未来很有帮助, 4分类: 完全不同意=1, ..., 完全同意=4	3.16	0.85
英语看法	英语对未来很有帮助, 4分类: 完全不同意=1, ..., 完全同意=4	3.17	0.92
家庭特征			
家庭经济状况	5分类: 非常困难=1, ..., 很富裕=5	2.81	0.59
兄弟姐妹	兄弟姐妹个数, 个	0.91	0.86
母亲教育	母亲受教育程度, 年	8.70	3.08
父亲教育	父亲受教育程度, 年	9.42	2.74
班级特征			
班级地位	班级在年级中的地位, 5分类: 最差=1, ..., 最好=5	3.20	0.96
班级规模	班级人数, 个	39.53	11.85
学校性质	公立学校=1, 非公立学校=0	95.79%	
语文教师			
性别	男=1, 女=0	29.14%	
受教育程度			
大学专科	是=1, 否=0	11.23%	
大学本科(成人)	是=1, 否=0	37.75%	
大学本科(正规)	是=1, 否=0	46.10%	
研究生及以上	是=1, 否=0	4.92%	
教学经验	从教年数, 年	14.53	9.14
年龄	岁	36.66	8.12
数学教师			
性别	男=1, 女=0	46.13%	
受教育程度			
大学专科	是=1, 否=0	18.13%	
大学本科(成人)	是=1, 否=0	38.69%	
大学本科(正规)	是=1, 否=0	42.02%	
研究生及以上	是=1, 否=0	1.17%	
教学经验	从教年数, 年	17.69	9.52
年龄	岁	39.41	8.37
英语教师			
性别	男=1, 女=0	13.79%	
受教育程度			
大学专科	是=1, 否=0	11.30%	
大学本科(成人)	是=1, 否=0	37.42%	

大学本科（正规）	是=1，否=0	48.82%	
研究生及以上	是=1，否=0	2.46%	
教学经验	从教年数，年	14.12	9.01
年龄	岁	36.05	7.60

附表6 留守儿童同伴效应回归结果：SUR模型

	语文	数学	英语
留守儿童比重	-1.5975** (-2.4788)	0.2712 (0.3345)	-1.2196* (-1.8629)
个体特征			
性别	-0.5043*** (-11.3309)	-0.2912*** (-6.6523)	-0.4685*** (-11.6139)
年龄	-0.0532** (-2.2007)	-0.0625** (-2.5188)	-0.0349 (-1.3512)
农业户口	0.0579 (1.4457)	0.0175 (0.4276)	-0.0187 (-0.4730)
本地户口	0.1477*** (2.7510)	0.0979** (2.1764)	0.0671* (1.6989)
健康	-0.0194 (-1.2632)	-0.0157 (-0.8352)	-0.0418** (-2.4134)
认知能力	0.3973*** (16.0296)	0.4634*** (13.4018)	0.4131*** (14.3906)
信心	0.0542** (2.1993)	0.0399* (1.6916)	0.0654*** (3.2265)
自评语文/数学/英语成绩	0.2004*** (9.8796)	0.2270*** (15.2589)	0.2301*** (13.6504)
语文/数学/英语看法	0.1032*** (5.4650)	0.1001*** (5.0474)	0.0872*** (6.3885)
家庭特征			
家庭经济状况	-0.0294 (-0.8448)	-0.0573* (-1.8487)	-0.0259 (-0.8438)
兄弟姐妹	-0.0326 (-1.4368)	-0.0502** (-2.0071)	-0.0629*** (-2.6881)
母亲教育	0.0001 (0.0117)	0.0018 (0.2680)	0.0002 (0.0317)
父亲教育	0.0079 (0.8638)	0.0022 (0.2719)	0.0119 (1.6069)
班级特征			

班级地位	0.1215** (2.5144)	0.0851** (2.0591)	0.2448*** (7.6776)
班级规模	0.0050 (0.4500)	0.0160* (1.6814)	0.0068 (0.7806)
学校性质	0.5690** (2.2936)	0.5239*** (3.9158)	0.5668*** (5.7661)
语文/数学/英语教师特征			
性别	-0.2161*** (-2.8225)	-0.1626*** (-2.7577)	-0.4567*** (-5.2526)
受教育程度（参照组：研究生及以上）			
大学本科（正规）	-0.3394** (-2.3128)	-0.1368** (-1.9975)	0.0109 (0.1349)
大学本科（成人）	-0.4109*** (-2.6347)	-0.1443* (-1.7593)	-0.2237** (-2.4236)
大学专科	-0.5128*** (-3.1886)	0.5575*** (4.1589)	-0.4665*** (-3.5580)
教学经验	-0.0309 (-1.4938)	0.0067 (1.3671)	0.0042 (0.6621)
年龄	0.0271 (1.1437)	-0.0037 (-0.5143)	-0.0372*** (-5.1542)
Atanhrho	0.5954*** (20.7538)	0.6982*** (24.4320)	0.7286*** (22.9740)
学校×年级固定效应	是	是	是
样本量	2684		
Log likelihood	-8368.8017		

注：标准误为以班级为聚类变量的聚类稳健标准误。下同。

附表 7 留守儿童同伴效应回归结果：不同留守儿童定义（SUR 模型）

	语文	数学	英语
留守儿童比重	-0.8085* (-1.7435)	-0.0648 (-0.1072)	-1.3112** (-2.4419)
控制变量	是	是	是
学校×年级固定效应	是	是	是
样本量	2684		
Log likelihood	-8367.3207		

附表 8 留守儿童同伴效应反事实检验结果：SUR 模型

	语文	数学	英语
留守儿童比重	-0.1055 (-0.3651)	0.0334 (0.1295)	-0.1563 (-0.4988)
样本量	2684		
Log likelihood	-8367.5658		
留守儿童比重	-0.2184 (-0.7177)	-0.1318 (-0.4801)	-0.1769 (-0.5282)
留守儿童家庭经济状况	0.0323 (0.7042)	0.1035 (1.3820)	0.0452 (1.2960)
留守儿童父母受教育程度	0.0017 (0.1228)	0.0068 (0.5431)	0.0136 (1.0526)
样本量	2625		
Log likelihood	-8182.3581		

附表 9 留守儿童同伴效应回归结果：剔除退学、转入和转出的样本（SUR 模型）

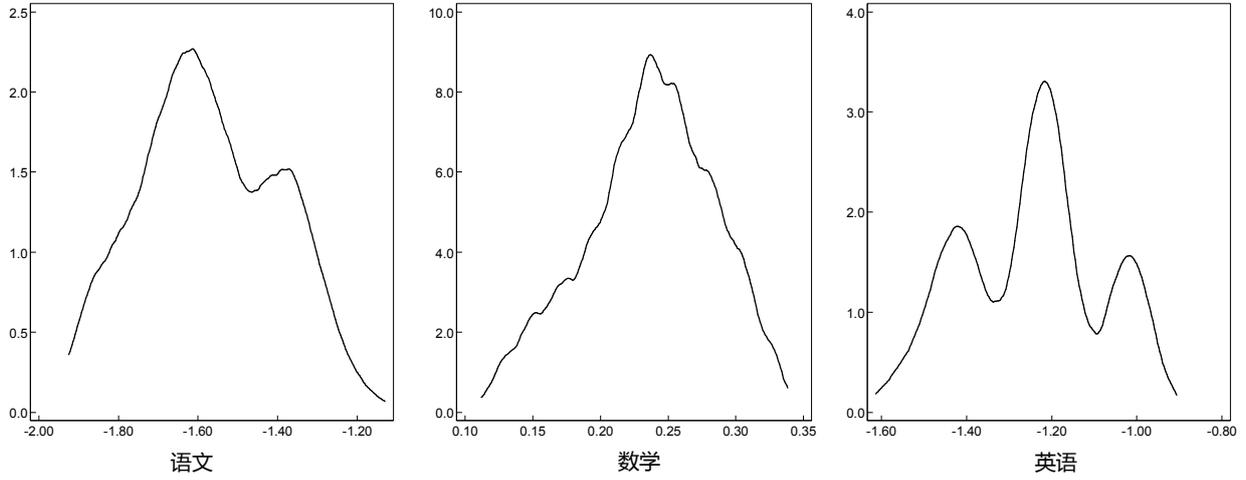
	语文	数学	英语
留守儿童比重	-1.8710** (-2.2525)	-0.0231 (-0.0345)	-1.9341* (-1.9389)
控制变量	是	是	是
学校×年级固定效应	是	是	是
样本量	1916		
Log likelihood	-5804.1446		

附表 10 留守儿童同伴效应回归结果：基于八年级学生数据（SUR 模型）

	语文	数学	英语
七年级留守儿童比重	-2.3800* (-1.6779)	0.6928 (0.5539)	-3.2434*** (-2.5878)
七年级留守儿童成绩	-0.0398 (-0.5089)	-0.0378 (-0.4092)	-0.1760*** (-2.9940)
样本量	1430		

Log likelihood

-4317.6415



附图 2 基于缩减样本的留守儿童同伴效应估计值分布状况

附表 11 留守儿童和不同家庭特征同伴的学习成绩

	留守儿童	父母受教育程 度<9年	父母受教育程 度≥9年	低收入家庭	中高收入家庭
语文	69.8843 (9.8332)	70.0929 (9.7731)	70.2836 (9.9248)	60.6862 (9.1174)	74.5175 (6.4612)
数学	70.0487 (9.8615)	70.1667 (9.8399)	70.2230 (9.9905)	60.6068 (7.9873)	74.6748 (7.0905)
英语	70.0218 (9.5480)	69.9474 (9.8990)	70.4881 (9.7808)	60.1434 (7.1684)	74.7425 (7.1332)

注：括号内为标准差。

参考文献

- [1] Bietenbeck J. The Long-Term Impacts of Low-Achieving Childhood Peers: Evidence from Project STAR[J]. Journal of the European Economic Association, 2019, 18(1): 392-426.
- [2] Guryan J, Kroft K, Notowidigdo M J. Peer Effects in The Workplace: Evidence from Random Groupings in Professional Golf Tournaments[J].

American Economic Journal: Applied Economics, 2009, 1(4): 34-68.

- [3] Ohinata A, Van Ours J C. How Immigrant Children Affect the Academic Achievement of Native Dutch Children[J]. The Economic Journal, 2013(123): F308-F331.