

农村中小学教育中寄宿制影响的 因果推断研究*

乔天宇 狄 雷

提 要：本文使用工具变量回归的方法，解决最小二乘回归在估计寄宿的影响时可能存在的内生性偏误问题，实证估计了寄宿对农村学生学习成绩的效应，从而考察寄宿对农村中小学教育的影响。研究表明，在校寄宿显著地提高了农村中小学生的学习成绩；对农村小学生和初中生群体分别加以考察，发现在校寄宿可能更显著地影响了农村小学生的学习成绩，而对农村初中生来说，寄宿并没有表现出显著的影响；与数学成绩相比，寄宿对语文成绩的效应更加突出。

关键词：寄宿制 农村学生 学习成绩 工具变量回归 因果效应推断

一、研究问题与文献回顾

为了更加合理地配置农村教育资源，从2001年开始，我国在农村地区大力推行和实施“撤点并校”政策。由于学校被裁撤，许多农村学生上学距离变远，往返家和学校之间的时间增加，这使他们不得不选择在学校寄宿。同时，随着我国现代化进程的发展，农村成年劳动力大规模向城市流动，引致了备受社会公众关注的留守儿童问题。由于日常缺少父母的看护，许多处于义务教育阶段的农村留守儿童不得不寄宿于学校。在这些需求驱动下，我国许多地方都在努力地发展农村寄宿制学校。教育部公布的数据显示，2011年底全国义务教育阶段学校寄宿生共有3276.51万人，占义务教育阶段在校生总数的比例为21.85%；其中，小学寄宿生1080.78万

* 本论文使用数据部分来自北京大学“985”项目资助、北京大学中国社会科学调查中心执行的中国家庭追踪调查。感谢北京大学社会学系许琪博士在写作过程中与作者的讨论，以及审稿人对本文提出的修改建议。

人,所占比例为 10.89%;初中寄宿生 2195.73 万人,所占比例为 43.34%^①。寄宿对于学生教育,尤其是对农村中小学教育来说,究竟是好事还是坏事?寄宿制作为一种社会设置在发挥其正功能的同时也必然会存在着一些反功能(Merton, 1968),寄宿对学生教育的影响也存在正反两方面的效应。从优势方面来说,寄宿制学校为学生提供了一个安全且稳定的课后成长环境(中央教育科学研究所课题组, 2008),相对延长了学生接受学校教育的时间,具有较强的可控制性等(浙江省磐安县实验小学课题组, 1998);从劣势方面来说,寄宿生可能会缺少父母亲的关爱,承受更大的心理压力。沙韦连(Schaverien, 2011)研究发现,早年的寄宿经历会增加遭遇意外伤害的概率,会对其之后的生活造成持续性的影响,他将这种影响称为“寄宿综合征(Boarding School Syndrome)”。另外,寄宿学校基础设施不足、师资力量较弱、学校管理较差等问题都可能会对学生的发展产生一定的消极影响(张茂聪、冯永刚, 2009;谢治菊、刘洋, 2012)。

寄宿的优势明显还是劣势更明显,或者说寄宿是有利于还是不利于农村的中小学教育,这一直是研究者争论的焦点。随着微观调查数据的丰富,可以通过实证研究评估农村中小学教育中寄宿的效应。学生的学业发展是考察教育成效的关键所在,而学习成绩是最为直接的衡量指标。因此,可以通过考察寄宿对农村学生学习成绩的效应,来评估寄宿对农村学生学业发展的影响,进而厘清寄宿在农村中小学教育中发挥的作用。

目前国内对农村寄宿学生学习成绩的研究较少。杜屏等(2010)在对西部五省区农村小学(四年级)寄宿生学习成绩的研究中发现,寄宿生的数学成绩显著地高于总体的平均水平,而语文成绩与总体平均水平相比并没有显著差异;胡伶和万恒(2012)通过调查发现寄宿生自我报告(self-reporting)的学习成绩相对保守;卢珂和林育红(2010)的研究发现寄宿生的成绩略低于非寄宿生,家校距离、家庭经济负担、学校配套设施不健全、学校心理疏导缺失等因素对学生学习成绩都可能产生负向影响。另外,钟奕(2011)和谢治菊(2013)等进行的个案研究也得出了不同的结论。现有少数几篇文献都是在某所学校或者某个地区开展的调查研究,所使用的样本数据不具备全国层面的代表性,所得结论并不能对我国农村学龄儿童

① 中华人民共和国教育部公布的《2011年全国教育事业发展统计公报》,详见:http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/moe_633/201208/141305.html。

的情况进行有效的推断。

另外,虽然已有研究试图去评估寄宿对农村中小学生学习成绩的影响,但鲜有研究去识别寄宿对农村学生学习成绩的因果效应。学生是否寄宿的选择极有可能是内生的,这是因为寄宿于学校对于学生(或对于学生的家庭)来说本身是一个有选择性的决策,具有某些特征的学生会更加倾向于选择寄宿。例如家庭经济状况可能会影响寄宿的决策,如果学生的家庭经济条件足够好,有能力支付学生上学的交通费用,甚或家长到学校附近陪读的成本,则这些学生选择寄宿于学校的可能性较小。同时,家庭的经济状况也会影响学生的学习成绩(Coleman, 1966)。上述文献也很可能存在遗漏变量的问题。例如,张林秀等(2009)的研究表明,农村学生是否寄宿于学校会影响学生的身体发育和健康状况,而这些生理因素同样会影响学生的学习成绩。但是一般的社会调查却很少测量诸如反映生理因素的指标,这也会导致分析所得的效应估计结果偏误。如果对这些潜在的内生性问题不加以考虑和控制,则很有可能得到误导性的结论。

本文试图通过实证研究估计寄宿对农村学生学习成绩的效应,来考察寄宿对农村中小学教育的影响。与既有研究不同的是,本文在考察寄宿的影响时,尝试着解决参数估计中可能存在的内生性偏误问题。同时,研究这一问题具有一定的政策意义。在多大程度与规模上建设和发展寄宿制学校,需要首先对寄宿发挥的实际作用有所了解,本研究的结果可以为类似的教育政策制定和改进提供一定的现实依据。

二、数据、变量与方法

(一) 数据

本文使用2010年中国家庭动态跟踪调查中少年儿童样本数据。中国家庭动态跟踪调查(以下简称CFPS)是由北京大学中国社会科学调查中心发起的全国性的社会调查项目,2010年实施的是基线调查,其以家庭户为调查的目标单元,采用分层多阶段PPS的抽样方法抽取村委会或居委会,末端使用简单随机抽样的方法抽取被调查的家庭户,并对抽中的家庭户中的所有成员进行访问,获得一个覆盖了全国25个省、市、自治区的代表性样本。在接受调查的14960个家庭户中,共有8990

名0-15岁的儿童回答了少儿问卷。由于本文的研究对象是农村中小學生，因此从该调查数据中选取目前居住在农村并且正在就读小学或初中的儿童作为分析对象。经筛选，得到一个包含1881名农村在读学生的样本。

(二) 模型与变量

本研究设定的回归模型的基本形式如下：

$$y_i = \beta_0 + \beta_1 x_{0i} + \gamma X_i + \lambda \alpha_i + \varepsilon_i \quad (1)$$

本文要考察的问题是寄宿是否有利于农村中小學生学习成绩的提高。无疑，想要讨论的因变量 y 是学习成绩，那么如何测量学习成绩是首先要解决的问题。CFPS(2010)中对所有调查对象都进行了标准化字词能力和数学能力测试，笔者拟采用这两项测试的得分对学习成绩进行测量，其中，用标准化字词能力测试的得分来衡量语文成绩，用数学能力测试的得分来衡量数学成绩。本文关注的核心自变量 x_0 为是否在学校寄宿，将在学校寄宿的农村学生视为干预组，不在学校寄宿的农村学生则为控制组。式(1)中的 X 是一组控制变量，包括个人特征变量——性别、年龄和教育阶段(小学为参照组)，家庭社会经济地位变量——父亲受教育年数、母亲受教育年数和家庭人均年收入^①，以及学生上学使用的交通方式和在上学路上花费的时间。

以往相关的研究表明，我国各地区的教育政策、资源投入以及教学情况差异极大，这会导致地区间中小學生学习成绩的差异(曾满超、丁延庆，2003；胡咏梅、卢珂，2010)。因此，本文将对地区层面的变异性予以控制，式(1)中的 α_i 即是对省份固定效应的控制。

(三) 工具变量回归

本文所关注的核心自变量是否在学校寄宿并非一个外生随机的干预，如果采用普通最小二乘法(OLS)进行估计，得到的效应结果可能会存在非常大的内生性偏误。所以，本文采用工具变量回归的方法(Instrumental Variable Method，下文中将

① CFPS(2010)公布的少儿问卷数据库中，其中一些基本信息有两套变量，一套是少儿自答的，另一套是受访儿童的家长回答的，这里的控制变量(家庭经济地位等)和核心自变量(是否寄宿)均统一使用家长回答的信息。

工具变量简称为 IV)，使用两阶段最小二乘法（2SLS）来估计寄宿对农村中小学生学习成绩的影响，以解决模型识别中可能会存在的内生性偏误问题。工具变量回归的基本形式如下：

第一阶段回归：

$$x_{0i} = \delta_0 + \delta_1 X_i + \delta_2 \alpha_i + \theta Z_i + r_{0i} \quad (2)$$

第二阶段回归：

$$y_i = \beta_0 + \beta_1 \hat{x}_{0i} + \gamma X_i + \lambda \alpha_i + \varepsilon_i \quad (3)$$

在式（2）和式（3）中， x_0 是本文关注的核心自变量， Z_i 则是选取的工具变量（IV），用第一阶段回归方程（2）中所得的 x_0 的拟合值 \hat{x}_0 替代方程（1）中的 x_0 ，得到第二阶段的回归方程（3），对其中 β_1 的估计即为内生自变量 x_0 的因果效应。

本研究选取的 IV 是农村学生家庭所在村庄的人口规模，该变量可以直接从 CFPS（2010）社区问卷的调查数据中获得。由于一个好的 IV 估计需要具备两个条件，一方面 IV 不能够与原方程（1）的残差项相关，另一方面 IV 必须与内生的自变量相关，即方程（2）中的 θ 必须显著地异于零。本文选取上述 IV 也是尽量地遵循这两个原则：一方面，村庄人口规模因素看上去与学生的学习成绩并没有直接的联系^①；另一方面，它与学生是否在学校寄宿有着一定的相关性。这是因为“撤点并校”政策的实施，改变了原来“村村办校”的局面，农村地区原有的大量中小学被裁撤，将教育资源集中于集镇或者在人口相对集中的地点进行办学（杨兰，2010）。在这种情况下，人口相对集中、规模相对较大的村庄更可能会成为“撤点并校”中教育资源的集中地，家住在这类村庄的学生在学校寄宿的可能性会比较小；相反，那些远离集镇中心且人口规模较小的村庄不太可能成为农村学校所在地，家住在这类村庄的学生则更可能会在学校寄宿。因此，如果这样一种分析逻辑成立，那么学生家庭所在村庄的人口规模与他们在学校寄宿的可能性之间应呈负相关关系。在这些假设之中，第一个是不能够直接地予以验证的，不过在一定的条件下，上述 IV 或许和学生的学习成绩还有一定的联系，要考虑它们之间可能存在一个共同影响因素——经济发展水平——家庭居住于经济条件更好社区的学生，其成绩可能会更好，而同时农村社区的经济发展水平也会与该村村委会到集镇的距离以

① 关于这一点，并非没有异议，后文中还将予以详细讨论。

及该村的人口规模相关 (Skinner, 1964)。关于这一点, 后文中还会做详细的讨论。而对于第二个假设, 即所选择的 IV 与内生自变量之间相关, 则可以通过第一阶段回归的结果来进行检验。表 1 是对本文中用到的所有变量的统计描述和操作化说明。

表 1 变量统计描述和操作化说明

变量	均值	标准差	最小值	最大值	操作化与变量类型
因变量①					
字词得分	20.699	7.383	0	34	语文成绩, 连续变量
字词得分的对数	2.984	0.509	0	3.555	
数学得分	10.648	4.512	0	24	数学成绩, 连续变量
数学得分的对数	2.353	0.512	0	3.218	
自变量					
是否寄宿 (寄宿 = 1)	0.302	0.459			二分变量
控制变量					
性别 (男 = 1)	0.494	0.500			二分变量
年龄	12.514	1.731	10	15	连续变量
教育阶段					
小学	0.667	0.471			二分变量
初中	0.333	0.471			
年级	5.607	1.848			
父亲受教育年数	4.068	2.579	1	14	连续变量
母亲受教育年数	2.948	2.255	1	13	连续变量
上学使用的交通方式					
步行	0.665	0.472			分类变量
自行车	0.161	0.368			
机动车	0.136	0.343			
其他	0.038	0.192			
上学路上花费的时间	23.804	23.838	1	180	单位为分钟, 连续变量
家庭人均收入的对数	8.204	0.871	3.239	11.705	家庭年总收入 (元) / 家庭规模, 连续变量
社区人均收入的对数	7.565	0.866	3.980	9.473	被调查村的年人均收入 (元), 连续变量
工具变量					
村庄的人口规模	21.083	16.931	1.7	130	单位为百人, 连续变量

① 这里分别对字词能力测试和数学能力测试的原始得分做了自然对数化的处理, 目的是在模型分析中方便对效应值进行解读, 以及对各模型之间的效应大小做出比较。

三、农村寄宿生的学习成绩：IV 估计的结果

(一) 寄宿生的学习成绩更好吗？

表 2 是分别对寄宿和不寄宿的农村中小学生的标准化字词能力测试得分和数学能力测试得分的描述统计分析结果，该表的最后一列给出了独立样本 T 检验的结果。

如果对所有的农村中小学生进行考察，可以发现，无论是字词测试还是数学测试，在学校寄宿的农村学生的成绩得分都要显著地高于那些不在学校寄宿的农村学生 ($P < 0.01$)。而如果将小学生和初中生分开考察的话，对于农村小学生来说，寄宿生与非寄宿生之间的差别很明显，同样无论是字词测试还是数学测试，在校寄宿的农村学生的得分会更高；而对于农村的初中生来说，寄宿生与非寄宿生均没有表现出显著的差异 ($P > 0.05$)。

表 2 寄宿生与非寄宿生标准化字词测试与数学测试的得分均值

		寄宿	不寄宿	双尾检验显著性
全样本	字词得分	23.808 (0.259)	19.354 (0.206)	$P < 0.001$
	数学得分	13.111 (0.162)	9.583 (0.120)	$P < 0.001$
	样本量	568	1313	
小学	字词得分	20.271 (0.461)	17.938 (0.223)	$P < 0.001$
	数学得分	9.402 (0.248)	8.219 (0.109)	$P < 0.001$
	样本量	199	1056	
初中	字词得分	25.715 (0.263)	25.175 (0.323)	$P = 0.193$
	数学得分	15.111 (0.115)	15.187 (0.160)	$P = 0.694$
	样本量	369	257	

注：括号内为标准差。

上述结果能否回答这样一个问题：更好的学习成绩真的是由在校寄宿带来的吗？根据前文中的假设和分析，由于是否在学校寄宿可能存在着内生性问题，因此需要

更加小心地求证寄宿对农村学生学习成绩的影响。

(二) 最小二乘回归与工具变量回归的结果

表 3 中的模型 I 报告的是控制了上面提到的个人特征以及其他相关变量之后的 OLS 估计结果。对于字词测试得分来说, 模型 Ia 估计出寄宿的效应大小为 0.057, 这意味着在控制了其他变量的条件下, 在学校寄宿的农村中小学生的字词测试得分比那些非寄宿生的得分高出了 5.9%^①, 并且这个差异在总体上是显著的 ($P < 0.05$); 而对于数学测试得分来说, 在控制了其他变量之后, 是否寄宿并没有表现出显著的影响。模型 II 是在模型 I 的基础上控制了省份固定效应后的结果。在数学测试得分上, 寄宿生与非寄宿生之间仍然没有表现出显著的差异, 而寄宿生与非寄宿生的字词测试得分差异明显增大(效应值变大且显著性水平提高)并且对省份虚拟变量的联合检验统计显著 ($P = 0.011$)^②。这个结果一方面说明, 我国中小学生的学习成绩在地区间存在着一定的差异, 也从侧面印证了既有研究的结论; 另一方面, 在校寄宿的农村学生的学习成绩确实要好于那些并不在校寄宿的农村学生, 且更加突出地表现在语文成绩上。同时, 寄宿生与非寄宿生学习成绩的差异并不大, 即便是在控制了省份固定效应之后, 农村寄宿生的字词测试得分也仅比非寄宿生高出 7.4%^③, 而在数学成绩上却都没有表现出显著的差异。

表 3 寄宿对农村中小学生学习成绩影响的因果推断: OLS 和 IV 估计的结果 (N = 1881)

	模型 Ia	模型 Ib	模型 IIa	模型 IIb	模型 IIIa	模型 IIIb	IV 一阶段估计
	(OLS)	(OLS)	(OLS)	(OLS)	(IV)	(IV)	
	字词得分	数学得分	字词得分	数学得分	字词得分	数学得分	
是否寄宿	0.057 ** (0.026)	0.031 (0.023)	0.072 *** (0.027)	0.031 (0.025)	0.669 ** (0.267)	0.468 ** (0.236)	
性别 (男 = 1)	-0.095 *** (0.020)	-0.005 (0.018)	-0.087 *** (0.019)	0.001 (0.018)	-0.100 *** (0.022)	-0.009 (0.020)	0.025 (0.016)
年龄	-0.015 (0.010)	-0.011 (0.009)	-0.017 (0.010)	-0.011 (0.009)	-0.020 * (0.012)	-0.014 (0.010)	0.006 (0.009)

① 该百分比是通过模型 Ia 估计出的寄宿效应 0.057 计算得到的, 计算方法为 $e^{0.057} - 1$ 。

② 限于篇幅, 文中未给出省份虚拟变量的估计系数, 有兴趣的读者可向笔者索要。

③ 该百分比是通过模型 IIa 估计出的寄宿效应 0.072 计算得到的, 计算方法为 $e^{0.072} - 1$ 。

续表

	模型 Ia	模型 Ib	模型 IIa	模型 IIb	模型 IIIa	模型 IIIb	IV 一阶段估计
	(OLS)	(OLS)	(OLS)	(OLS)	(IV)	(IV)	
	字词得分	数学得分	字词得分	数学得分	字词得分	数学得分	
初中	-0.161 *** (0.037)	0.114 *** (0.033)	-0.160 *** (0.036)	0.119 *** (0.033)	-0.320 *** (0.082)	0.002 (0.072)	0.269 *** (0.031)
年级	0.169 *** (0.012)	0.160 *** (0.011)	0.166 *** (0.012)	0.158 *** (0.011)	0.153 *** (0.014)	0.148 *** (0.013)	0.023 * (0.010)
父亲受教育年数	0.019 *** (0.004)	0.006 * (0.004)	0.015 *** (0.004)	0.006 (0.004)	0.012 *** (0.005)	0.003 (0.004)	0.005 (0.003)
母亲受教育年数	0.007 (0.005)	0.005 (0.004)	0.006 (0.005)	0.004 (0.005)	0.012 * (0.006)	0.008 (0.005)	-0.010 * (0.004)
上学使用的交通方式 (参照组: 步行)							
自行车	0.078 *** (0.029)	0.034 (0.026)	0.052 * (0.029)	0.026 (0.027)	0.058 * (0.033)	0.030 (0.029)	-0.017 (0.025)
机动车	0.051 (0.033)	0.047 (0.029)	0.031 (0.033)	0.041 (0.030)	-0.190 * (0.105)	-0.121 (0.093)	0.366 *** (0.027)
其他	0.116 ** (0.053)	0.093 ** (0.047)	0.065 (0.054)	0.087 * (0.050)	-0.158 (0.116)	-0.076 (0.103)	0.374 *** (0.046)
上学路上花费的时间	-0.024 (0.026)	-0.049 ** (0.023)	-0.018 (0.026)	-0.042 * (0.024)	-0.162 ** (0.070)	-0.147 ** (0.062)	0.238 *** (0.022)
家庭人均收入	0.029 ** (0.012)	0.022 ** (0.011)	0.019 (0.012)	0.020 * (0.011)	0.016 (0.013)	0.018 (0.012)	0.007 (0.010)
省份	—	—	控制	控制	控制	控制	控制
IV: 村庄的人口规模	—	—	—	—	—	—	-0.003 *** (0.001)
截距	1.959 *** (0.140)	1.333 *** (0.125)	2.072 *** (0.258)	1.513 *** (0.237)	2.276 *** (0.301)	1.662 *** (0.266)	-0.314 (0.221)
R ²	0.306	0.453	0.357	0.464	0.358	0.465	0.422

注: (1) * p < 0.1, ** p < 0.05, *** p < 0.01; (2) 括号内为稳健标准误。

表 3 的最后三列报告的是 IV 两阶段最小二乘 (2SLS) 估计的结果。首先, 从第一阶段回归的结果来看, 我们所选取的 IV 与内生解释变量 (是否寄宿) 是显著相关的 (P < 0.01, F 值为 23.784), 这在一定程度上排除了“弱工具变量”的风险 (伍德里奇, 2007: 502), 并且其参数估计的结果与假设的方向相一致, 即农村学生家庭所在村庄的人口规模越大, 其选择在学校寄宿的可能性则越低。这些都满足了 IV 估计的基本条件。其次, 模型 III 报告的是第二阶段回归的结果, 能够发现, 无论是对于字词测试得分还是数学测试得分, IV 估计结果与 OLS 估计的结果都有着很

大的区别：一方面估计效应值都有所增加，在校寄宿使得字词得分提高了 95.2%，使得数学成绩提高了 59.7%^①，另外与 OLS 估计结果不同之处是，寄宿对于字词成绩和数学成绩的效应均通过了统计显著性检验。

（三）讨论 1：距离与经济发展水平

上文讨论工具变量选取的时候，提到农村社区的经济发展水平与所使用的 IV 可能存在一定的相关性——经济发展水平越高的农村社区，其拥有的人口规模可能会相对更大，家住此类村庄的学生在学校寄宿的可能性会相对较小。也就是说，学生家庭所在农村社区的经济发展水平可能是本文所选工具变量与学生学习成绩之间的一个共同影响因素。如果这一假设成立，IV 与式（1）的残差项不相关的假设就不能够得到满足，极可能破坏 IV 估计的结果，同样带来估计的偏误。在这一部分，我们将试图排除这种可能的干扰，检验上面所结果的稳健性。

本文以社区人均纯收入的对数来测度农村社区的经济发展水平，与工具变量一样，社区人均纯收入也可以从 CFPS2010 社区问卷的调查数据中获得，对该变量的描述详见表 1。由第一阶段回归结果发现，在控制了社区人均收入变量之后，对 IV 回归系数的估计结果和检验的显著性水平都没有发生太大的变化，社区人均收入的效应为负，也符合我们此前所做出的推断。表 4 中的模型 IVa 和模型 IVb 报告的是控制了农村社会经济发展水平之后第二阶段回归的结果，与模型 III 的结果相比，能够发现寄宿所带来效应的大小也基本没变，对字词得分的效应稳定在了 0.67 左右，而对数学得分的效应略有增加，但变化幅度也不是很大，并且均在 0.05 的水平上统计显著，能够说明本节第二部分中得到的分析结果在一定的条件下是比较稳健的。

（四）讨论 2：基于家长报告的学习成绩

由于 CFPS2010 的少儿问卷中还向受访学生的家长询问了该学生在学校的学习情况，为了进一步检验上述研究结果的稳健性，本文又使用了学生家长报告的学习成绩情况作为因变量，其余部分的模型设定和分析策略基本保持不变，以检验寄宿所带来的效应是否发生了变化。

① 有关效应值的大小及工具变量的有效性问题，本文在结论一节会做进一步的说明。

表 4 控制了社区经济发展水平后寄宿对农村中小学生学习成绩影响的因果推断 (N = 1881)

	模型 IIIa 字词得分	模型 IVa 字词得分	模型 IIIb 数学得分	模型 IVb 数学得分
是否寄宿	0.669 ** (0.267)	0.674 ** (0.271)	0.468 ** (0.236)	0.491 ** (0.241)
其他控制变量	控制	控制	控制	控制
省份	控制	控制	控制	控制
社区人均纯 收入的对数		0.005 (0.015)		0.023 * (0.013)
截距	2.276 *** (0.301)	2.257 *** (0.302)	1.662 *** (0.266)	1.582 *** (0.269)
R ²	0.358	0.358	0.465	0.465

注: (1) * p < 0.1, ** p < 0.05, *** p < 0.01; (2) 括号内为稳健标准误。

对于家长报告的学习成绩变量, 有两点需要说明: 一方面, 与前面使用的标准化的字词能力测试和数学能力测试得分不同, 它属于对学生学习成绩的相对测量^①。由于我国各地区间的教育政策和学生的学习状况有相当大的差异, 而在现实生活中, 人们更加倾向于将自身的状况与周围人的状况进行比较, 对于学习成绩则更是如此, 学生的成绩只有放在班级内部、学校内部或者地区内部比较可能更有意义。因此, 与使用标准化测试得分相比, 使用对学习成绩的相对测量方法可能更具有现实意义。同时, 此种测量为本文提供了另一种可能的稳健性检验方法, 如果以其为因变量, 进行检验的结果并没有发生太大的变化, 则说明此前的研究结论在一定程度上更具稳健性。另一方面, 家长报告的学习成绩变量从理论上来说属于有序的分类变量, 但为了方便进行 IV 估计, 这里使用的是线性概率模型 (Linear Probability Model), 将家长报告的四个等级分别赋值为优 = 4、良 = 3、中 = 2、差 = 1, 并对其进行回归, 而非处理有序分类因变量的 Logit 模型 (Ordered Logit Model) (伍德里奇, 2007: 283 - 284)。

表 5 的左半部分是对家长报告学习成绩的描述统计结果, 右半部分是以家长报告的学习成绩为因变量来设定模型所估计的在校寄宿带来效应的结果。该结果同样

① 这两个家长报告的变量原始访题编号为 F501 和 F502, 提问方式为“就您所知, 孩子上学期平时的语文成绩 (数学成绩) 如何”。

显示,在校寄宿能够显著地提高学生的语文成绩和数学成绩;将其对语文成绩和数学成绩的效应进行对比,结果显示在校寄宿对语文成绩的效应要大于其对数学成绩的效应。上述结果与此前以标准化的测试得分作为因变量的研究结果并没有太大变化,这进一步说明了此前所得的分析结果是比较稳健的。

表 5 对家长报告学习成绩的描述以及对寄宿带来效应的 IV 估计

	家长报告的 语文成绩	家长报告的 数学成绩	自变量	模型 Va 语文	模型 Vb 数学
优 = 4	0.119 (0.324)	0.173 (0.378)	是否寄宿	1.758 *** (0.606)	1.158 ** (0.594)
良 = 3	0.374 (0.484)	0.335 (0.472)	其他控制变量①	控制	控制
			省份	控制	控制
中 = 2	0.299 (0.458)	0.287 (0.452)	截距	3.795 *** (0.694)	3.317 *** (0.680)
差 = 1	0.208 (0.406)	0.206 (0.405)	N②	1844	
			R ²	0.147	0.106

注: (1) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; (2) 左半部分描述统计报告的是均值, 括号中为标准差, 右半部分括号中为稳健标准误。

四、寄宿对谁更有好处?

在讨论了寄宿影响农村中小学生学习成绩的因果效应之后, 本文更进一步分别考察寄宿对农村小学生和初中生学习成绩的影响。在上述分析的基础上, 本文将农村中小学生学习成绩样本分成了农村小学生和农村初中生两个子样本, 在分析策略上依旧沿用上文中的模型设定方法, 对它们分别进行两阶段最小二乘估计。表 6 是参数估计的结果。

① 这里的其他控制变流量包括模型 IV 中的所有其他控制变量和社区人均纯收入的对数。

② 在符合条件的 1881 名受访农村在校中小学生学习成绩中, 有 37 名学生的家长对他们平时在校的学习成绩表示并不知道, 将这 37 名学生作为有缺失数据的观测案例, 在这部分的分析中并不予以考虑, 故进入分析模型的样本量是 1844。

表 6 寄宿对农村中小学生学习成绩影响：分别考察小学生和初中生的情况

	小学生		初中生	
	字词得分	数学得分	字词得分	数学得分
是否寄宿	0.788 (0.568)	0.774 * (0.452)	0.206 (0.216)	-0.138 (0.132)
其他控制变量	控制	控制	控制	控制
省份	控制	控制	控制	控制
截距	2.050 *** (0.315)	1.461 *** (0.361)	2.791 *** (0.183)	2.105 *** (0.136)
N	1255	626		
R ²	0.143	0.021	0.081	0.103

注：(1) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; (2) 括号中为稳健标准误。

表 6 的结果比较出乎意料。可以发现，对农村小学生来说，在校寄宿对数学测试得分呈现出了显著的正效应 ($P < 0.01$)，且字词得分和数学得分二者上的效应值均要大于表 4 模型 IV 中相对应的平均效应；而对农村初中生来说，在校寄宿对两项测试的得分的影响均不显著，并且效应大小也远低于农村小学生样本，这与第三部分中通过描述性分析所得的结果是一致的。这在一定程度上可以说明，寄宿对农村小学生来说会更有好处。如何理解这一现象，笔者认为可能存在着两种解释。一方面，这可能与青少年的成长阶段有一定的关系，由于小学生的年龄较小，自制能力相对缺乏，尤其很多的农村小学生又属于留守儿童，他们身边缺乏父母的管制与约束，而在学校寄宿使得他们能够获得较长时间的看护，且能够受到学校规范的约束。有调查发现，学生在学校的绝大部分时间都是在上课或者上自习（姚姿如，2011），这些对于学习成绩的提高可能都更具积极作用；另一方面，还可能与不同学习阶段的学习任务和压力相关，如初中生的课后作业等学习任务会更多，寄宿的影响就不那么明显了。

五、结 论

本文使用工具变量回归的方法研究了在学校寄宿对农村中小学生学习成绩的影响，尝试着解决实证估计中是否在学校寄宿这一关键自变量的内生性问题，从而对农村中小学教育中寄宿所发挥的影响做出因果推断。根据以上两部分的分析结果，

基本上可以对本文所提出的研究问题做一个简单的回答。本文研究所呈现的基本结论是,对于农村的中小学生来说,寄宿的确能够影响其学习成绩,在学校寄宿会使学习成绩得到显著提高,文中得到的效应估计值可以认为是寄宿对农村中小学生学习成绩影响的因果效应。但是同时,也还需要对这一结论做更具体的分析和考察。

一是寄宿对农村中小学生语文成绩和数学成绩的效应是有区别的,本文的研究结果显示,寄宿对提高语文成绩的作用更大;

二是分析结果显示寄宿对农村小学生和农村初中生这两个群体的影响模式存在一定的区别,它显著地影响了农村小学生群体数学成绩,而对农村初中生群体的学习成绩并没有表现出显著的影响。

本文使用工具变量回归的方法,试图解决内生性问题,分析结果发现,IV估计的效应值大于OLS估计的效应值,这说明使用OLS回归得到的寄宿对农村中小学生学习成绩的影响可能会低估在校寄宿实际上所带来的效应。我们还对所选择的工具变量与学习成绩二者间共同的影响因素做了简单的讨论,排除了农村社区经济发展水平这个共同影响可能对IV估计造成的干扰;同时,使用家长报告的学习成绩变量作为因变量,重新估计与检验之前的模型,分析结果表明IV估计的效应是比较稳健的。

根据上述研究结论基本可以认为,寄宿在一定程度上能够提高农村中小学生的学习成绩,对农村中小学生的学业发展是有益的,尤其对农村小学生来说,影响更加突出。总的来说,在发展农村中小学教育和促进农村中小学生成长的问题上,寄宿还是发挥了较为积极的作用,这一结论也将为我国政府目前在农村地区推行促进寄宿制学校建设的政策提供较有利的现实依据。

本文也存在一定的不足:首先在工具变量的选择上,本文假定了学生家庭所在村的人口规模与学生的学习成绩之间除了农村社区经济发展水平之外不再存在其他的共同影响因素,当然这是一个非常强的假定条件,而IV估计结果的有效性则是要基于这个强假定,之所以认为当前的结论合理,也是基于认为这个假定是基本合理的;其次,文中使用IV估计方法计算得到寄宿的效应结果要远大于使用传统OLS方法估计所得的结果,这是否也指向了IV有效性?关于这一点,本文尚不能够给出比较合理的解释。所以,关于IV的有效性,并不能完全排除存在其他的可能性,若存在更加合理的逻辑或假设,能够揭示出二者之间存在着其他的共同影响因素,便需要对IV的有效性进行重新评估;最后,本文发现寄宿可能对农村小学生更有好处,而对原因只给出了一些简单的假设,并没有进一步地验证,留待以后研究讨论。

参考文献:

- 杜屏、赵汝英、赵德威, 2010,《西北五省区农村小学寄宿生学业成绩与学校适应性研究》,《教育学报》第6期。
- 胡伶、万恒, 2012,《农村寄宿制学生社会情感学习能力调查》,《中国教育学刊》第9期。
- 胡咏梅、卢珂, 2010,《教育资源投入对学生学业成绩的影响力评价——基于西部地区基础教育发展项目的研究》,《教育学报》第6期。
- 卢珂、林育红, 2010,《农村学校布局调整对学生成绩的影响——基于两水平增值模型的分析》,《清华大学教育评论》第6期。
- 伍德里奇, 2007,《计量经济学导论》,费剑平译,北京:中国人民大学出版社。
- 谢治菊, 2013,《教育均衡视域下农村寄宿制学校建设探析——基于访谈记录的分析》,《教育财会研究》第4期
- 谢治菊、刘洋, 2012,《边远贫困山区农村寄宿制学校建设研究》,《中国教育学刊》第8期。
- 杨兰, 2010,《农村中小学撤点并校政策回顾与反思》,载于《中国教育发展报告2010》,北京:社会科学文献出版社。
- 姚姿如, 2011,《丰富农村寄宿制学校生活的思考》,《东北师大学报(哲学社会科学版)》第3期。
- 曾满超、丁延庆, 2003,《中国义务教育财政面临的挑战与教育转移支付》,《北京大学教育评论》第1期。
- 张林秀、刘承芳、罗仁福、史耀疆、岳爱、Scott Rozelle, 2009,《农村小学生的贫血状况及其对身体发育和学业表现的影响》,2009年中国教育经济学会年会会议论文。
- 张茂聪、冯永刚, 2009,《公平与均衡视域下义务教育资源有效配置的研究——以山东省农村中学生进城就读现象为例》,《教育研究》第12期。
- 浙江省磐安县实验小学课题组, 1998,《农村小学寄宿制管理模式综合实验研究报告》,《教育研究》第8期。
- 中央教育科学研究所课题组, 2008,《贫困地区农村寄宿制学校学生课余生活管理研究》,《教育研究》第4期。
- 钟奕, 2011,《农村寄宿制学生的学习现状研究》,《现代教育科学:小学教师》第5期。
- Coleman, James S. 1966, *Equality of Education Opportunity*. Washington, DC: Government Printing Office.
- Merton, Robert K. 1968, *Social Theory and Social Structure*. New York: The Free Press.
- Schaverien, Joy. 2011, "Boarding School Syndrome: Broken attachments a hidden trauma." *British Journal of Psychotherapy* 27 (2).
- Skinner, G. William. 1964, "Marketing and Social Structure in Rural China, Parts I, II, and III." *Journal of Asian Studies* 24 (1).

作者单位: 北京大学社会学系(乔天宇)

山东师范大学公共管理学院(狄雷)

责任编辑: 张巍巍

market of processing industry; thirdly, the broker's social relations have great impact on the acquisition of orders; fourthly, the strength of the broker's social relations plays a very important role in contract signing and the structure of the order.

The Causal Inference of Boarding's Effect in Rural Primary and Secondary School Education
..... *Qiao Tianyu & Di Lei* 138

Abstract: This article is trying to reduce the endogenous bias in estimating the effect of boarding due to OLS by using the method of instrumental variable regression. The effect of boarding on the academic performance of rural students is thus estimated, revealing the influence of boarding on the rural primary and secondary education. As the results show, boarding in school could significantly improve the academic performance of rural primary and secondary school students. When considering the students of primary and secondary schools in rural area separately, students of primary school are found to be more probable to be affected by boarding, while such effect on secondary school students is not significant. When considering the academic performance by subject, the effect of boarding on improving Language is greater than on Mathematics.

The Research on the New Type of the Urban Home Church in City B: The case study of university student Christians
..... *Liu Zheng'ai* 153

Abstract: This paper studies the developmental characteristics of university Christian groups in city B through the study on the process of establishment and development of a new type of house churches and Christian fellowships. It finds that these urban home churches share one common characteristic; they are initiated by one particular person. One possibility is that the Christian organized bible study groups at home or college campus. And gradually, these groups developed to bigger scale churches. Another possibility is that a Christian who left the church for some reasons organizes a home church or fellowship, and develops the home church into larger church. This article discusses the mechanisms of how these house churches recruit their followers and indicates that interpersonal relationship plays a key role in these recruiting channels.

Inauguration: A study of the local political order in Ch'ing Dynasty
..... *Wang Shaochen* 170

Abstract: In Ch'ing Dynasty, when a government official was assigned to a new post in an unfamiliar administrative area, he would encounter two institutional challenges. First, due to the restriction of the