

接受知青占人口比例对当地教育水平的影响(河北省)

一、小组分工(姓名 学号 工作量占比)

崔士强	PB22151743	30%
张博伦	PB22151731	40%
郑中锐	PB22151751	30%

二、背景介绍:

上世纪四十到七十年代的教育,经历了知青上山下乡的变化,在此期间,千万知青离开家乡来到周边的城镇和乡村插队,知识分子的到来也为当地教育带来了不少改变。

本实验重点关注各城市接受知青比例对于当地教育水平的影响,并希望从诸多指标的变化中总结出二者的相关性。

三、数据统计与总述:

1 数据组成

以有无接受知青作为自变量,设置了对照组和实验组(通过前十年和后十年作为组别界限),分别统计了在乡村和城镇中的受教育状况、人口比例情况、城镇化进展等数据,以此分析知青接受比例对当地教育水平的影响。

2 分析思路

通过纵向对比十年变化和横向对比城乡差距,以不同城镇学历人口比例等诸多数据,作为教育水平的指标,再与知青接受比例相比较,最后得出相关性结论。在本文中,我们只选取了几个与教育水平相关的指标作为研究对象。

四、数据分析：

基于人口普查数据的概要统计(全国以及河北省)

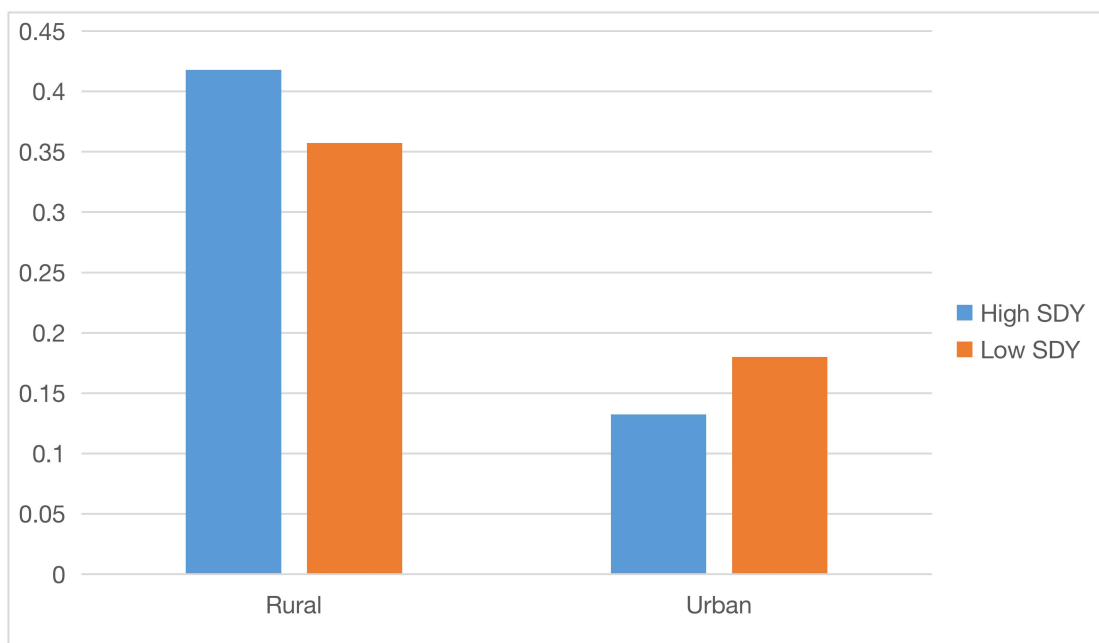
Nationwide				
Cohort	Control group (1946-1955)		Treatment Group (1956-1969)	
Hukou	Rural (1)	Urban (2)	Rural (3)	Urban (4)
Years of education	5.099 (1.503)	8.823 (1.844)	6.957 (1.508)	10.222 (1.556)
Complete primary school	0.575 (0.226)	0.893 (0.171)	0.763 (0.197)	0.953 (0.126)
Complete junior high	0.192 (0.122)	0.659 (0.257)	0.432 (0.190)	0.860 (0.193)
Male=1	0.507	0.592	0.505	0.567
Han ethnic=1	0.924	0.930	0.914	0.925
Observations	960,127	148,244	1,815,752	271,603

Note: Standard deviations are in parentheses.

Provincial				
Cohort	Control group (1946-1955)		Treatment Group (1956-1969)	
Hukou	Rural (1)	Urban (2)	Rural (3)	Urban (4)
Years of education	5.699 (0.942)	9.188 (1.926)	7.686 (0.714)	10.845 (1.757)
Complete primary school	0.749 (0.136)	0.948 (0.144)	0.898 (0.071)	0.977 (0.091)
Complete junior high	0.269 (0.091)	0.731 (0.251)	0.578 (0.116)	0.888 (0.227)
Male=1	0.497	0.654	0.501	0.583
Han ethnic=1	0.960	0.923	0.948	0.894
Observations	59,278	4,414	103,860	6,561

Note: Standard deviations are in parentheses.

1. 受教育时长与知青比例的关系

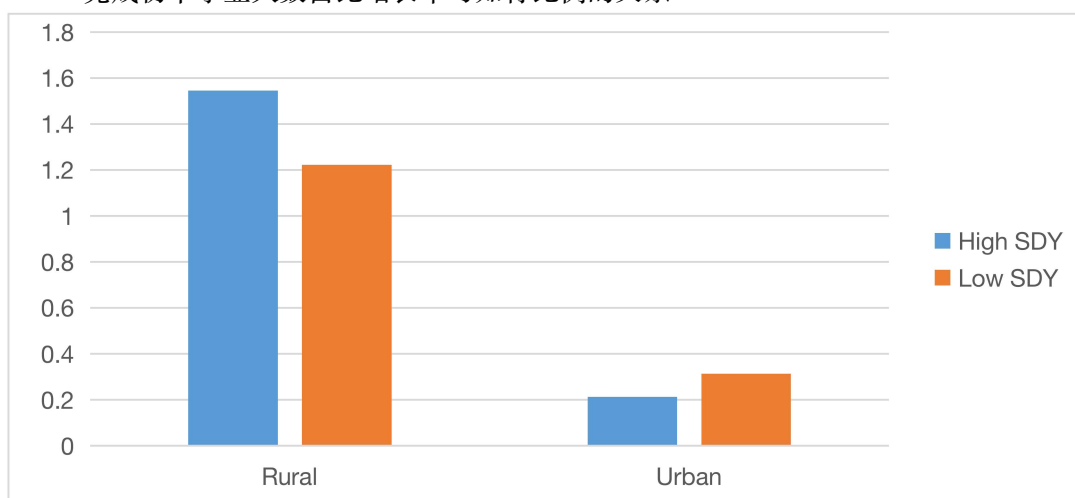


图表解释：纵轴数值表示实验组相应对照组的增长率的平均值

由图表可知：在乡村，高知青比例下的受教育时长增长的幅度更大，但是在城市则正好相反，反倒是地知青比例的地区的受教育时长提高更多。这一结论与我们平常认为的知青促进教育的看法相悖，看似反直觉，但是实际上仔细分析一下数据来源就可知，本实验对于城市知青高低水平仅仅通过接受知青这一指标来评判，但是忽视了城市也同时是知青输出的主要地区，并且输出数量会随城市的教育水平的提高而提高，所以很可能出现了教育水平高的地区反而知青比例较低的情况。

排除掉上述无关因素的影响，我们仅从乡村地区（几乎不存在知青输出的影响）的数据来分析，便可以大致得到，知青比例的提高可以显著地提高当地受教育年限，从而提高当地教育水平。

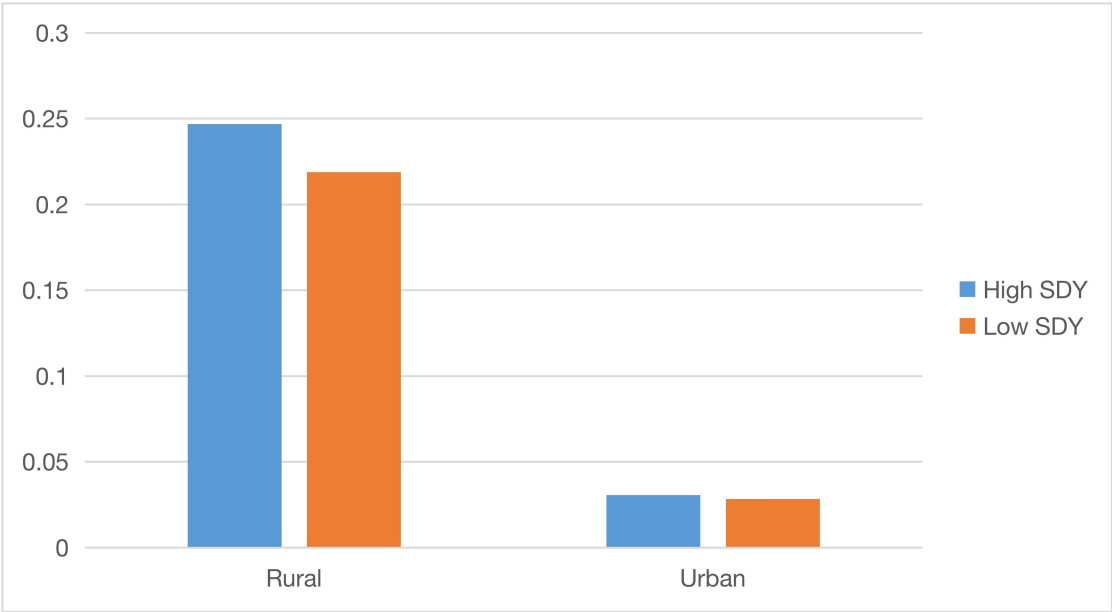
2. 完成初中学业人数占比增长率与知青比例的关系



图表解释：纵轴数值表示实验组相应对照组的增长率的平均值

由图表可知，关于完成初中学业人数占比，实验组都相比对照组有所增长，其中乡村的平均增长率显然更大，同时这个指标也出现了低知青城市增长率高于高知青城市增长率的情况，推测和上一个指标的原因相同。此外，我们重点关注增长率的数值大小，不难发现，乡村地区的完成初中学业人数占比提高了一倍多，而在本来教育水平就相对发达的城市，初中学业人数上升的则没有那么多，由此可见，知青上山下乡会显著提高当地初中学业水平的人数，从而提高教育水平。

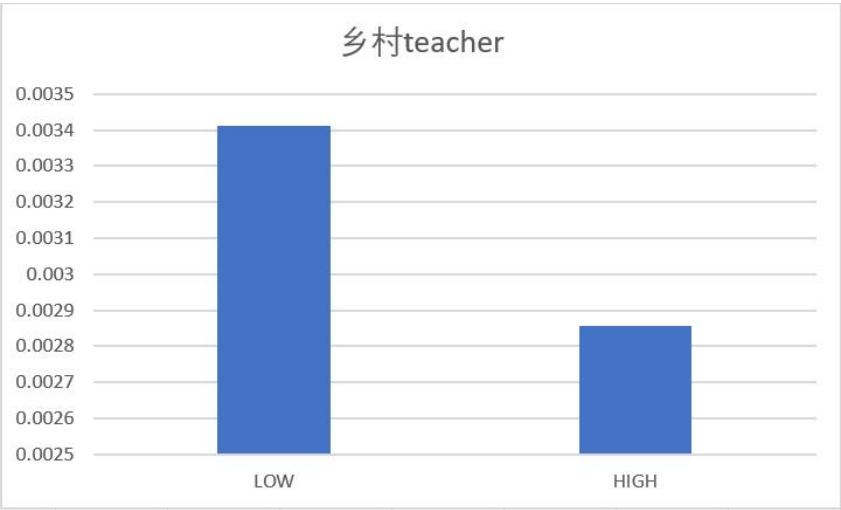
3.完成小学学业人数占比增长率与知青比例的关系

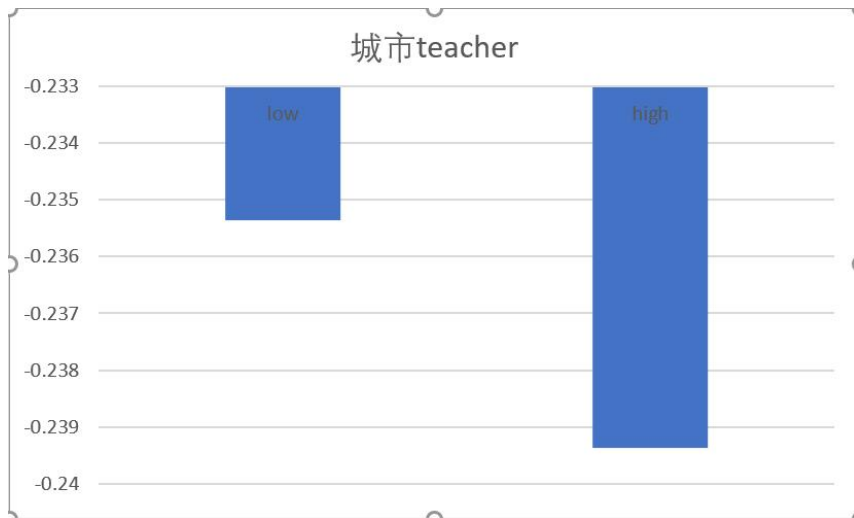


图表解释：纵轴数值表示实验组相比对照组的增长率的平均值

由图表可知，实验组都相比对照组有所增长，规律也和之前所述大致相同，但和上一个初中的指标相比，知青对完成小学学业的人数的提高并不明显，尽管乡村仍然远超城市增长率，但是数值很小，说明在知青上山下乡前后，河北省的小学教育已经较为普及，所以知青的到来并没有带来太大影响。

4.教师占比增长率与知青比例的关系





图表解释：纵轴数值表示实验组相对对照组的增长率的平均值

由图表可知，知青的到来对乡村教师比例的影响微乎其微，而且高低知青接受比例也对教师比例影响很小；而知青上山下乡则降低了城市教师的比例，至于为什么高接受知青的城市的教师比例反而下降的更多，笔者也没有思考出合理解释，可能是数据来源对高知青城市的评判出了问题。

5.其他指标

笔者从仅有的数据中，并没有分析出性别比和汉族比例对于当地教育水平的影响。

五、分析数据来源以及统计周期：

总数据源于文章 Chen, Fan, Gu and Zhou (2020) "Arrival of Young Talent: The Send-Down Movement and Rural Education in China ", American Economic Review , 110(11): 3393-3430。

数据提供者将数据集中于城市水平，本文分析者则重点关注河北省 1946 ~ 1969 年的数据，希望从单个省份的数据特点总结出一般性规律。

六、对当前数据的反思：

1. 本实验设计的对照组和实验组位于两个不同的十年，无法验证排除时间变化对于教育发展的影响，即无法做到完全的控制变量。
2. 部分实验组和对照组的数据不对应，即出现了数据丢失的情况。
3. 有不少数据指标都没有显著差异，可以看出数据指标的选择还有待精进，可以适当去除一些与教育无关的指标。
4. 考虑到派出知青的地区绝大多数是城市，但是在评定城市知青数量的高低水平时，却只考虑了接受数量，没有统计知青输出数量，导致城与城、城与乡之间会有不小知青水平差别，甚至出现相反的数据结论。由此可知，有些数据指标的评判还不尽合理，仍有待优化。