说明文档

（一）文件说明

本论文使用的数据代码压缩包命名为“2023-01026+数据”，解压之后里面有“2023-01026+**程序代码**”、“2023-01026+**日志文件**”和“2023-01026+**说明文档**”3个部分。

（1）“2023-01026+**程序代码**”里面包括4个文件，分别是“主要回归结果”、“健康机制检验”、“家庭经济机制检验”和“父母对子女照料陪伴机制检验”文件：

“主要回归结果”是文章主要结果运行文件，主要生成表1描述性统计、表2、表3、图1、图2和在线附录稳健性检验结果。

“健康机制检验”是文章健康机制检验的运行文件，主要生成表4结果。

“家庭经济机制检验”是文章家庭经济机制的运行文件，主要生成表5结果。

“父母对子女照料陪伴机制检验” 是文章父母对子女照料陪伴机制的运行文件，主要生成表6结果。

（2）“**2023-01026+日志文件**”中的“表1”是表1描述性统计；“表2”、“表3” 、“表4”、“表5”、“表6”分别是正文表2-表6的运行结果。“图1”、“图2” 分别是正文图1-图2的运行结果。“表I1”、“表II1”、“表III1”、“表IV1”、“表V1”、“表VI1”、“表VII1”分别是在线附录表I1、表II1、表III1、表IV1、表V1、表VI1、表VII1的运行结果。

（3）“**2023-01026+说明文档**”即是本说明文档。

（二）数据来源

本文使用的数据包括三部分：

（1）北京大学中国社会科学调查中心（ISSS）实施并发布的中国家庭追踪调查（China Family Panel Studies，CFPS）数据。CFPS自2010年开展基线调查以来每隔两年展开一轮追踪调查，目前已经进行了六轮追踪调查。尤为重要的是，CFPS采用概率比例规模抽样方法（PPS）从全国25个省（市/自治区）收集个人、家庭和社区层面的数据，其样本具有良好的全国代表性，能够有效反映我国居民教育、健康、劳动和收入等状态的变迁，是当前我国可供研究使用最权威的微观个体数据之一，被学术界所广泛使用。出于如下两点原因，我们使用CFPS 2018年数据进行研究。其一，由于本文主要关注1998年“两控区”政策的长期人力资本收益，与2018年相隔20年，政策实施前后出生的个体按照年龄计算已经基本完成高中学历教育，部分正在读大学，这为我们检验政策对个体教育成就的影响提供了机会。 其二，2019年末新冠疫情席卷全球，对个体健康造成了巨大影响，因此如果使用2020年的CFPS数据可能会导致我们的识别模型存在较多的混淆因素，不利于我们的因果识别。

（2）国务院网站公布的“两控区”实施地点名单。1998年1月，国务院印发了《国务院关于酸雨控制区和二氧化硫污染控制区有关问题的批复》（国函〔1998〕5号），其附件包含了全部酸雨控制区和二氧化硫污染控制区的名单，我们据此作为“两控区”的实施范围。

（3）在机制检验中我们使用中国家庭健康与营养调查（China Health and Nutrition Survey，CHNS）的1991、1993、1997、2000和2004年共5轮调查数据进行分析。CHNS始于1989年，是由北卡罗来纳大学人口研究中心、美国国家营养与食物安全研究所和中国疾病与预防控制中心合作开展的调查项目。CHNS采用多阶段整群抽样的方法，调查范围涵盖了我国经济发达、中等和较差地区，调查内容覆盖了健康、营养、就业、收入等多项主题，是当前可用于考察我国上世纪个体和家庭变迁最权威的面板追踪数据之一。

（三）Dofile里面使用的指标说明

highschool：是否有高中学历或者正在上高中；

university：是否有大学学历或者正在上大学；

treat：出生区县是否属于“两控区”实施范围，是取1，否则取0；

post：“两控区”实施时个体是否在-1到5岁，是取1，否则取0；

han：是否汉族，是取1，否则取0；

sex：男性取1，女性取0；

hukou\_3：3岁时户口，非农取1，农业取0；

countyid：个体出生区县代码；

birthyear：个体出生年份；

provcd：个体出生省份代码；

cmonth：调查月份；

interviewerid18：访问员代码；

mathtest：数学测试得分；

wordtest：字词测试得分；

cognitive：数学测试和字词测试得分均值

ill\_child：过去四周是否生病；

lnheight：身高对数；

height\_score：HAZ得分；

haz\_dummy：根据HAZ得分判断的是否健康，是取1，否则取0；

age\_mean：父母平均年龄；

yedu\_f：父亲教育年限；

yedu\_m：目前教育年限；

divorce：父母是否离婚，是取1，否则取0；

hhsize：家庭规模；

lnhhinc\_pc：家庭人均收入对数；

citycode：个体出生区县代码（CHNS）

year：调查年份

job：是否有工作，有取1，否则取0；

lnwage：工资对数；

lnhhinc\_pc：家庭人均收入对数；

medical2：家庭医疗支出；

nutrition：家庭营养摄入；

dianqi：家庭耐用电器数量

disorder: 是否被医生诊断患有内分泌紊乱、精神系统疾病、智障和神经系统疾病任意一种，是取1，否则取0；

zhaogu：上周是否照顾陪伴子女，是取1，否则取0

zhaoguhour：上周平均每天照顾陪伴子女小时数