

附录

A. 微观理论补充

考虑一个具有个体异质性的人力资本投资模型。高中毕业生 i 需要最大化整个生命周期的效用,不妨将其简化为两期。在第零期,其决定是否通过上大学进行人力资本积累;在第一期,则基于个人人力资本积累水平获取收入。则高中毕业生 i 的决策目标为:

$$\begin{aligned} & \max_{(I_i, c_{0,i}, c_{1,i})} U(c_{0,i}) + \beta U(c_{1,i}) \\ \text{s.t. } & c_{0,i} + \frac{1}{1+r} c_{1,i} \leq w_{0,i}(EHC_i(I_i)) + \frac{1}{1+r} w_{1,i}(EHC_i(I_i)) \end{aligned}$$

其中 $U(\cdot)$ 为单期效用函数, β 为主观贴现因子。对于 $t=0,1$, $c_{t,i}$ 分别表示个体消费额, r 为市场借贷利率; $w_{t,i}(EHC_i(I_i))$ 为与预期人力资本水平 $EHC_i(I_i)$ 有关的工资,不妨设 $w_{t,i}(\cdot)$ 为单调递增函数。 I_i 为个体是否读大学的决策变量,即个体决定读大学时 $I_i=1$,否则 $I_i=0$ 。由于学校教育并不是进行人力资本积累的唯一方式,因此读大学就要放弃相应年限的工作经验以及可能存在的在职培训。那么基于个体特征的预期人力资本积累水平由下式决定:

$$EHC_i(I_i) = I_i \cdot f_e(u_i) + (1 - I_i) \cdot g(u_i) + \varepsilon(u_i) \quad (A1)$$

其中 u_i 为个体特征, $f_e(u_i)$ 表示个体 i 读大学对人力资本积累的事前预期贡献, $g(u_i)$ 表示不读大学而直接参加工作对人力资本积累的贡献, $\varepsilon(u_i)$ 表示个体高中毕业时所具有的人力资本存量。个体将根据个人特征 u_i 进行人力资本投资决策。易知,个体 i 决定读大学的条件为

$$w_{0,i}(EHC_i(1)) + \frac{1}{1+r} w_{1,i}(EHC_i(1)) \geq w_{0,i}(EHC_i(0)) + \frac{1}{1+r} w_{1,i}(EHC_i(0))$$

忽略大学期间的工资性收入¹,则个体 i 决定读大学的条件简化为

$$w_{1,i}(EHC_i(1)) \geq w_{1,i}(EHC_i(0))$$

则个体 i 的人力资本投资决策可由下式表示:

$$I_i = \begin{cases} 1, & f_e(u_i) \geq g(u_i) \\ 0, & f_e(u_i) < g(u_i) \end{cases} \quad (A2)$$

易知,基于个人效用最大化的决策结果(A2)同正文中基于收入最大化的决策结果相同。

B. 计量理论补充

为识别边际处理效应需要有额外的工具变量 Z ,则选择方程变为

$$D = 1\{p(R, Z) > V\} \quad (B1)$$

在式子(1)和(A1)下, Heckman and Vytlacil (2005)给出以下假设识别边际处理效应:

A1: 条件于 X , $p(R, Z)$ 为非退化随机变量;

A2: 条件于 X , (U_0, V) 和 (U_1, V) 独立于 Z ;

A3: V 在 Lebesgue 测度下绝对连续;

A4: $E|Y_0| < \infty$ 且 $E|Y_1| < \infty$;

A5: $1 > \Pr(D = 1|X) > 0$;

A6: $X_1 = X_0$ 几乎处处成立。

核心识别定理的证明如下。

定理 1: 在假设 2.1-2.2 下, 有

¹ 一方面, 国内大学生读大学期间的工作强度不大甚至没有工作, 因此其大学期间的工资性收入可以忽略不计; 此外, 未读大学的高中毕业生其毕业前四年的收入相对生命周期收入(以18岁毕业工作到60岁计, 其工作年限为42年)占比较低, 为简化分析将其忽略。

$$\frac{\partial E[Y|R=r_0^+, X]}{\partial p_+} = E[Y_1 - Y_0|V = p_+, R = r_0, X] \quad (B2)$$

证明: 为简化符号以下证明均条件于 X , 并略写 X 。

$$\begin{aligned} E[Y|R = r_0^+] &= E[(Y_1 - Y_0)D + Y_0|R = r_0^+] \\ &= E[(Y_1 - Y_0) \cdot 1\{P(R) > V\}|R = r_0^+] + E[Y_0|R = r_0^+] \\ &= E[(Y_1 - Y_0) \cdot 1\{P(R) > V\}|R = r_0^+] + E[Y_0|R = r_0] \\ &= \lim_{r \rightarrow r_0^+} E[(Y_1 - Y_0) \cdot 1\{P(r) > V\}|R = r] + E[Y_0|R = r_0] \\ &= \lim_{r \rightarrow r_0^+} \int_0^{p(r)} E[(Y_1 - Y_0)|R = r, V = v] dv + E[Y_0|R = r_0] \\ &= \int_0^{p_+} E[(Y_1 - Y_0)|R = r_0, V = v] dv + E[Y_0|R = r_0] \end{aligned}$$

$$\text{则 } \frac{\partial E[Y|R=r_0^+, X]}{\partial p_+} = E[Y_1 - Y_0|V = p_+, R = r_0, X]。$$

收入的方差反映了收入的离散程度, 可以用于度量收入分配的不平等程度, 即收入的方差越大, 意味着收入的离散程度越大, 收入分配的不平等程度也就越大。定理 2 识别了政策强度对收入方差的影响, 可用于评价扩招扩招政策对收入分配不平等的影响。给定录取率下, 若上述参数为负值, 则说明扩招政策降低了收入分配的不平等程度, 反之则说明扩招政策增加了收入分配的不平等程度。此外, 我们可以进一步识别出政策强度对收入分布的影响, 见定理 3。

定理 2: 在假设 2.1-2.2 下, 有

$$\begin{aligned} &\frac{\partial \text{Var}[Y|R = r_0^+, X]}{\partial p_+} \\ &= E[Y_1^2 - Y_0^2|V = p_+, R = r_0] - 2E[Y|R = r_0] \cdot E[Y_1 - Y_0|V = p_+, R = r_0] \end{aligned} \quad (B3)$$

证明: 以下均默认条件于 X , 已知

$$\text{Var}(Y) = E(Y^2) - (E(Y))^2$$

则

$$\frac{\partial \text{Var}(Y|R = r_0^+)}{\partial p_+} = \frac{\partial E(Y^2|R = r_0^+)}{\partial p_+} - 2E(Y|R = r_0^+) \frac{\partial E(Y|R = r_0^+)}{\partial p_+}$$

已知

$$\begin{aligned} E(Y^2|R = r_0^+) &= E((DY_1 + (1-D)Y_0)^2|R = r_0^+) \\ &= E(DY_1^2 + (1-D)Y_0^2 + 2D(1-D)Y_1Y_0|R = r_0^+) \\ &= E(DY_1^2 + (1-D)Y_0^2|R = r_0^+) \\ &= E(D(Y_1^2 - Y_0^2)|R = r_0^+) + E(Y_0^2|R = r_0^+) \end{aligned}$$

而

$$\begin{aligned} &E(D(Y_1^2 - Y_0^2)|R = r_0^+) \\ &= \lim_{r \rightarrow r_0^+} E((Y_1^2 - Y_0^2) \cdot 1\{P(r) > V\}|R = r) + E(Y_0^2|R = r_0) \\ &= \lim_{r \rightarrow r_0^+} \int_0^{p(r)} E(Y_1^2 - Y_0^2|R = r, V = v) dv + E(Y_0^2|R = r_0) \\ &= \int_0^{p_+} E(Y_1^2 - Y_0^2|R = r, V = v) dv + E(Y_0^2|R = r_0) \end{aligned}$$

因此

$$\frac{\partial E(Y^2|R=r_0^+)}{\partial p_+} = E(Y_1^2 - Y_0^2|R=r_0, V=p_+)$$

于是得证下式

$$\begin{aligned} & \frac{\partial \text{Var}(Y|R=r_0^+)}{\partial p_+} \\ &= E(Y_1^2 - Y_0^2|R=r_0, V=p_+) - 2 \cdot E(Y|X, R=r_0) \cdot E(Y_1 - Y_0|R=r_0, V=p_+) \end{aligned}$$

定理 3: 在假设 2.1-2.2 下, 有

$$\frac{\partial F_{Y|R,X,Z}(Y|R=r_0^+, X=x, Z=z)}{\partial p_+} = F_{Y_1|V,R,X}(y|p_+, r_0, x, z) - F_{Y_0|V,R,X}(y|p_+, r_0, x, z). \quad (B4)$$

证明: 以下均默认条件于 X, Z ,

$$\begin{aligned} F(Y|R=r_0^+) &= E(1\{Y < y\}|R=r_0^+) \\ &= E(1\{Y < y\}|R=r_0^+, D=1)p_+ + E(1\{Y < y\}|R=r_0^+, D=1)(1-p_+) \\ &= E(1\{Y_1 < y\}|R=r_0^+, D=1)p_+ + E(1\{Y_0 < y\}|R=r_0^+, D=1)(1-p_+) \\ &= E(1\{Y_1 < y\}|R=r_0^+, P(R) > V)p_+ + E(1\{Y_0 < y\}|R=r_0^+, P(R) < V)(1-p_+) \\ &= \lim_{r \rightarrow r_0^+} \frac{\int_0^{p(r)} E(1\{Y_1 < y\}|R=r, V=v) dv}{p(r)} p_+ \\ &+ \lim_{r \rightarrow r_0^+} \frac{\int_{p(r)}^1 E(1\{Y_1 < y\}|R=r, V=v) dv}{1-p(r)} (1-p_+) \\ &= \int_0^{p_+} E(1\{Y_1 < y\}|R=r, V=v) dv + \int_{p_+}^1 E(1\{Y_1 < y\}|R=r, V=v) dv \end{aligned}$$

上式等号两边同时对 p_+ 求导即可得式 (B2)。

C. 半参数模型的估计方法

若不对 $m_0(\cdot), m_1(\cdot)$ 的函数形式加以限制而采用非参数方法来估计, 则会面临维数灾难。尤其是在控制变量 X 的维数较高的情况下, 为得到有效的估计量所需数据量是实证研究所难以满足的。但 $m_0(\cdot), m_1(\cdot)$ 采用完全线性的形式则面临较大的模型误设的风险。为在模型设定的灵活性和实证研究的可行性之间取得平衡, 本文采用半参数模型。其具体设定如下:

$$Y_0 = \beta_0 X + f_0(R) + U_0; \quad (C1)$$

$$Y_1 = \beta_1 X + f_1(R) + U_1 \quad (C2)$$

本文将 $f_0(R), f_1(R)$ 设定为 R 的三次函数, 这一设定下, 目标参数政策的边际强度效应为:

$$\begin{aligned} MIE &= \frac{\partial E[Y|R=r_0^+, X=x]}{\partial p_+} \\ &= E[Y_1 - Y_0|V=p_+, R=r_0, X=x] \\ &= (\beta_1 - \beta_0)x + (f_1(r_0) - f_0(r_0)) + E[U_1 - U_0|V=p_+, R=r_0, X=x] \end{aligned} \quad (C3)$$

在 (U_0, V) 和 (U_1, V) 独立于 (R, X, Z) 的假定下,

$$\begin{aligned} & E[U_1 - U_0|V=p_+, R=r_0, X=x] \\ &= E[U_1|V=p_+] - E[U_0|V=p_+] \end{aligned}$$

由 Carneiro and Lee (2009) 可知,

$$E[U_1|V=p_+] = E[U_1|V=p_+, D=1] + p_+ \frac{\partial E[U_1|V=p_+, D=1]}{\partial p_+} \quad (C4)$$

$$E[U_0|V=p_+] = E[U_0|V=p_+, D=0] - (1-p_+) \frac{\partial E[U_0|V=p_+, D=0]}{\partial p_+} \quad (C5)$$

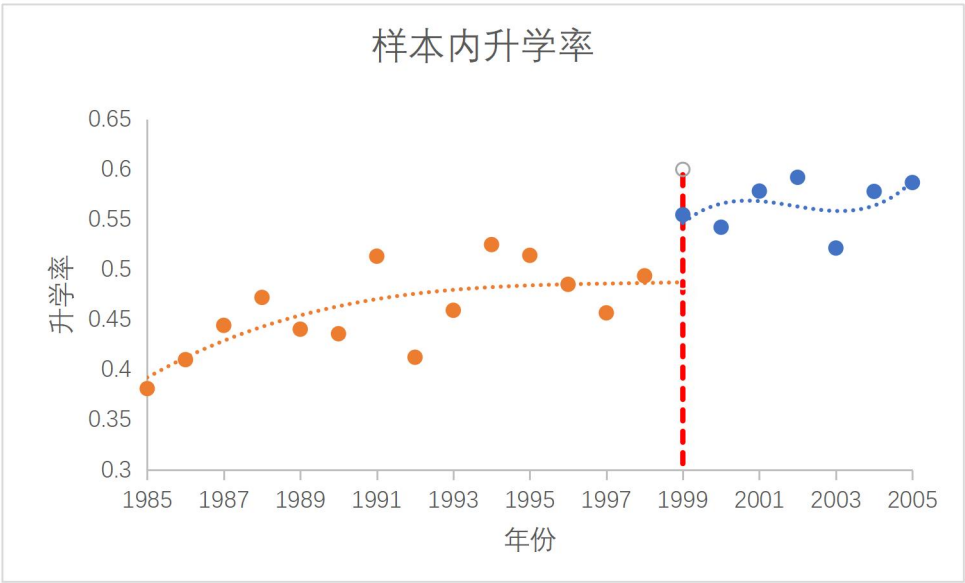
则可根据非参数方法估计 $E[U_1|V = p_+]$ 和 $E[U_0|V = p_+]$ ，而不需要担心控制变量过多所带来的维数灾难。基于以上讨论，本文建议三步法估计目标参数。第一步：通过参数或非参数方法估计出个体倾向得分 P ；第二步：根据已估出倾向得分使用 Robinson（1988）所提出方法估计 (β_0, β_1) ，以及 $f_0(R), f_1(R)$ 中 R 前的系数；第三步，根据第二步的估计结果计算残差 $(\widehat{U}_0, \widehat{U}_1)$ ，然后通过非参数方法估计 $E[U_1|V = p_+]$ 和 $E[U_0|V = p_+]$ 。

D. 数据的描述性统计

附表 D1 数据的描述性统计				
变量名	最小值	均值	最大值	标准差
对数工资	3. 81	10. 32	14. 40	0. 82
是否完成高等教育	0	0. 50	1	0. 50
参加高考年份	1985	1995. 46	2005	6. 23
是否为女性	0	0. 45	1	0. 50
农村户口	0	0. 31	1	0. 46
工作性质	0	0. 40	1	0. 49
已婚	0	0. 90	1	0. 30
是否为少数民族	0	0. 04	1	0. 20
是否为党员	0	0. 21	1	0. 41
是否健康	0	0. 88	1	0. 33
家庭资产（万元）	0	8. 64	362. 60	15. 70
父母受教育程度	0	0. 20	1	0. 40
兄弟姐妹个数	0	1. 53	9	1. 36
地区平均年收入	950	2464	9657	1541. 02
（元）				

样本内高等教育录取率变化趋势如下附图 D1。其中 1999 年后样本内录取率十分接近真实录取率，而 1998 年及之前样本内录取率高于实际录取率。分别以 3 次多项式拟合 1999 年之前和之后的高等教育录取率会发现在录取率在 1999 年出现明显跃升。不过跃升幅度只有 6%低于实际跃升幅度 21%。这种现象的主要原因是 1998 年及之前样本内录取率过高，这可能是样本内城镇户籍人口比例过高导致的²。断点回归设计以断点作为工具变量，因此样本内 98-99 年录取率跃升幅度不足的会导致断点回归结果出现偏误。而样本选择模型不需要以断点作为工具变量，因此更加契合数据集。

² 笔者对不同的微观调查数据进行处理，未能发现跃升幅度显著更大的数据集，因此最终选择 CHIP2013。



附图 D1 样本内升学率变化趋势图

E. Probit 模型的估计结果

由于本文控制变量和工具变量较多，因此最终选择 Probit 模型来估计个体倾向得分。另外，考虑到 1999 年前后录取率变化较大，因此建议使用两个模型分别估计 1985-1998 和 1999-2005 年个体倾向得分。所估计倾向得分的取值范围为[0.0347,0.9998]，其中位数为 0.4908，均值为 0.4992。参数的估计结果如下附表 E1。

附表 E1 Probit 模型的估计结果

变量	1998 年及之前	1999 年及之后
参加高考的年份	0.0216*** (0.0084)	0.0205 (0.0185)
兄弟姐妹个数	-0.0332* (0.0178)	-0.0981*** (0.0317)
当地平均工资	0.00025*** (0.00006)	0.000038 (0.000026)
女性	-0.0417 (0.0476)	0.2762*** (0.0614)
农村户口	-0.8286*** (0.0640)	-0.0491*** (0.0643)
在国有部门工作	0.6571*** (0.0499)	0.6767*** (0.0710)
已婚	-0.5466*** (0.1980)	-0.3328*** (0.0755)
少数民族	0.01586 (0.1257)	0.0440 (0.1529)
党员	0.6961*** (0.0572)	0.6736*** (0.1001)
身体健康	0.1100* (0.0658)	-0.2021* (0.1210)
资产	0.0080***	0.0109***

	(0. 0017)	(0. 0031)
父母教育	0. 2353***	0. 5304
	(0. 05712)	(0. 0887)
截距项	-0. 2301	0. 1548
	(0. 2517)	(0. 1574)
样本量	3488	2087
AIC	3914	2401

注：以 1999 年为基年对参加高考的年份进行处理，即 $R = \text{参加高考年份} - 1999$ ；*、**和***分别表示在 10%、5%和 1%的显著性水平下显著。

F. (β_0, β_1) 的半参数估计结果

为估计政策强度效应要首先估计出个体倾向得分以及不同学历水平下控制变量前的系数 (β_0, β_1) 。倾向得分可以衡量个体接受并完成高等教育的可能性，而 (β_0, β_1) 可以反映性别、户籍状况等控制变量如何影响个体收入。并且 β_0 与 β_1 之间的区别则可以反映不同学历水平下个人收入影响机制之间的差异。

下附表 F1 给出 (β_0, β_1) 的估计结果³，其中加入参加高考年份的三阶多项式控制时间趋势，并通过 Bootstrap 方法计算标准误。从附表 F1 中可以发现除了婚姻状况和民族状况对个体收入的影响不显著外，其余变量对收入的影响方向与学历无关。性别差异会对收入产生显著影响，其中女性的中学毕业生收入比男性低 32. 52%，对于大学毕业生这一数值为 25. 02%，这或许反映了劳动力市场中的性别歧视。此外，无论是否完成高等教育，农村户籍人口的收入均低于城镇户籍人口，中学毕业生的这一数值为 18. 02%，而大学毕业生的这一数值为 7. 49%。对于中学毕业生在国有部门工作将使其收入提升 7. 74%，但是否在国有部门工作对大学毕业生收入影响不显著。此外身体健康状况也会影响个体收入，而且身体健康对中学毕业生收入的影响也大于大学毕业生，这或许是因为中学毕业生更多的从事体力劳动，因此身体状况对收入的影响更大。总结附表 F1 的结果可以发现，个人的出生条件（例如性别、户籍状况）会对收入产生影响，但个体可以通过高等教育来减弱出生条件对收入的影响。高等教育对个体突破出生条件限制并实现更好发展是有积极意义的。

附表 F1 (β_0, β_1) 的半参数估计结果

变量	中学毕业对数收入	大学毕业对数收入
女性	-0. 3252*** (0. 0339)	-0. 2502*** (0. 0260)
农村户口	-0. 1802*** (0. 0357)	-0. 0749* (0. 0377)
国有部门工作	0. 0774** (0. 0370)	0. 0233 (0. 0291)
已婚	-0. 0627 (0. 0649)	0. 0220 (0. 0472)
少数民族	-0. 2070* (0. 1031)	0. 0122 (0. 057)
党员	-0. 0978* (0. 0542)	-0. 0009 (0. 0315)
身体健康	0. 1442*** (0. 0486)	0. 1095** (0. 0424)
资产	0. 0149***	0. 0078***

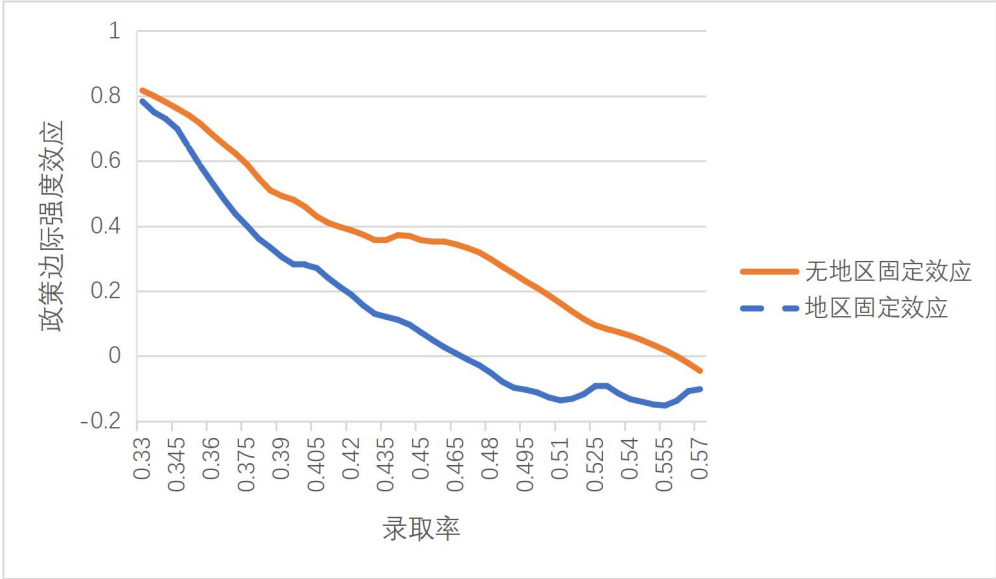
³ 条件于 $V = p_+$ 的 (β_0, β_1) 半参数估计结果会略微受 p_+ 取值的影响，附表 F1 的估计结果基于 $p_+ = 55.5\%$ 计算。

	(0.0027)	(0.0017)
父母初中毕业	0.0889*	0.1612***
	(0.0475)	(0.0281)
$f(R)$	控制	控制

注：*、**和***分别表示在 10%、5%和 1%的显著性水平下通过检验。

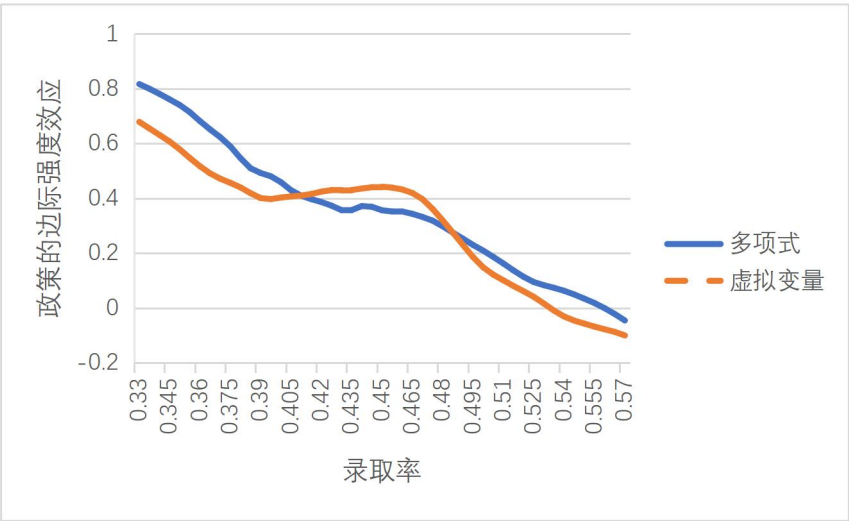
G. 稳健性分析补充

根据个体原始户口所在地区生成 7 个虚拟变量作为控制地区差异的变量，分别是华北地区（北京、天津、河北、山西和内蒙古）、东北地区（黑龙江、吉林、辽宁）、华东地区（上海、江苏、浙江、安徽、福建、江西、山东）、华中（河南、湖北、湖南）、华南地区（广东、广西、海南）、西南地区（重庆、四川、贵州、云南、西藏）、西北地区（陕西、甘肃、青海、宁夏、新疆）。结果如附图 G1 所示，无论是否加入地区固定效应来控制地区差异，扩招政策的边际效应均随着录取率的提高而下降。只不过控制地区差异后政策的边际强度效应下降得更快，在 47%的录取率下已经接近 0。



附图 G1 控制地区差异

不通过高考年限的多项式而是通过参加高考年份的虚拟变量来控制时间趋势的估计结果如附图 G2 所示。扩招政策的边际影响随着录取率的提高而下降的趋势依然保持不变。



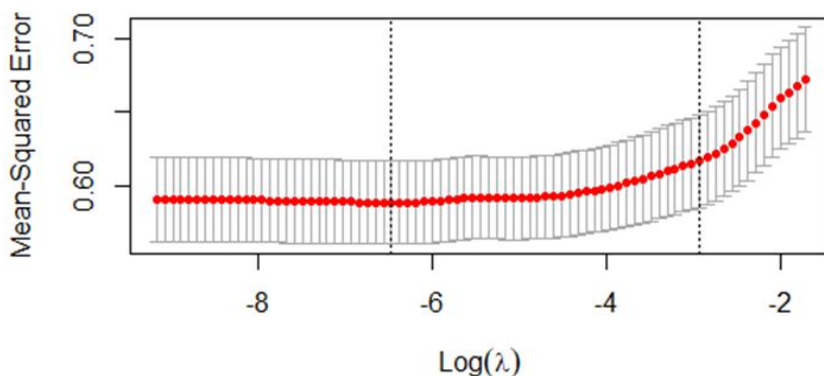
附图 G2 不同时间趋势控制方法

本文在上述估计扩招政策的边际强度效应的过程中采用了半参数设定,以在确保估计可行性的基础上尽可能地放松对模型设定的约束。事实上,如何在灵活的模型设定和简便可行的估计之间取得较好的权衡是实证研究者普遍较为关心的问题。随着机器学习等大数据方法的出现,为解决该类问题提供新的范式,本小节将基于此进行稳健性检验。

在上述估计过程中,本文设定控制变量 X 对隐含收入 (Y_0, Y_1) 只存在线性影响。为进一步放松模型设定,可以考虑加入 X 的交互项和高阶项。但直接增加交互项和高阶项会极大地增大多重共线性的风险。而且从 (β_0, β_1) 的半参数估计结果可以看到,部分控制变量的影响并不显著。因此不加选择地直接增加交互项和高阶项会导致模型中存在很多不显著的变量,这会对政策边际强度效应的估计结果产生不良影响。因此应该考虑在进行具体估计工作之前增加一个变量筛选过程。

在大数据时代,从众多变量中筛选出合适的变量是基于数据进行决策的重要问题,也是机器学习和人工智能研究中的热点问题。目前机器学习方法在经济学研究中主要应用于预测方面,而在因果推断方面应用较少。但相关研究正在快速推进,其中 Belloni *et al* (2012) 以及 Belloni *et al* (2014)⁴等创新性地将机器学习方法和因果推断相结合,提出首先通过 LASSO 或 post-LASSO 方法选择出最优的控制变量和工具变量组合,然后再进行正式的因果推断,并指出通过该方法可以实现高维数据因果推断。

本文借鉴 Belloni *et al* (2012) 的思想来进一步放松对模型设定的约束并进行稳健性检验。具体方法如下:首先构造出 X 的交互项、二次项和三次项;然后,通过 LASSO (Least Absolute Shrinkage and Selection Operator) 的方法从上述所构造变量以及 X 筛选出对 Y 具有非零影响的变量;最后,基于所筛选出来的控制变量采用半参数方法估计扩招政策的边际强度效应。显然惩罚项 λ 取不同的值会筛选出不同数目的变量,本文根据交叉验证法 (Cross-Validation) 计算出不同 λ 下的拟合值和真实值之间的均方误差。如附图 G3 所示,其中横坐标为 $\log(\lambda)$ 的取值,而纵坐标则为均方误差,中间的曲线则是交叉验证曲线以及上下标准差曲线。如下图所示,随着 λ 不断接近零,模型的拟合度逐渐提高,因此均方误差不断下降。但在 $\log(\lambda) \approx -3$ 时,继续减小 λ 已不足以使均方误差下降一个标准误, $\log(\lambda) \approx -6.5$ 时,均方误差达到最小值。为防止模型出现过多不显著的变量,本文最终设定 $\log(\lambda) \approx -3.2$,并筛选出 13 个对个体收入具有非零影响的控制变量⁵。



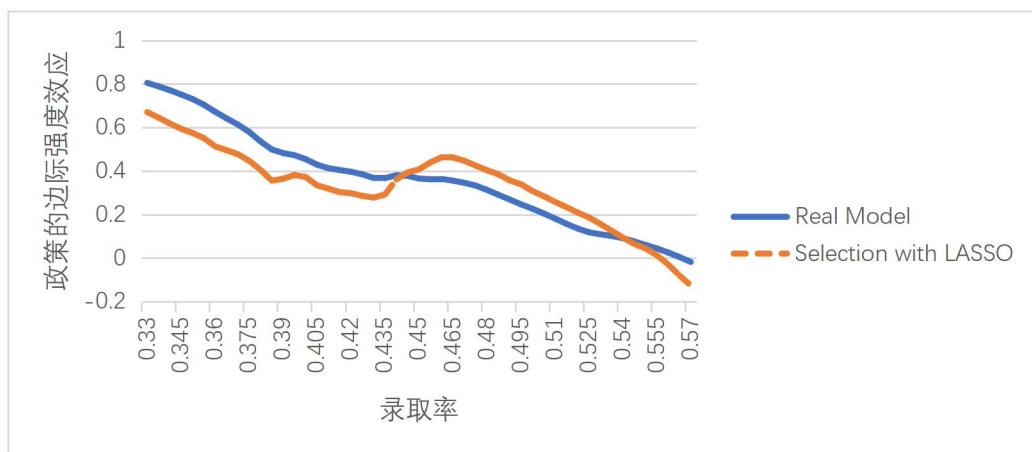
附图 G3 不同 λ 下的均方误差

基于以上方式筛选出的 13 个控制变量,再次估计 1999 年扩招政策的边际强度效应,其边际强度效应曲线如附图 G4 所示。其中实线为第四小节中的基准估计结果,虚线则为基于 LASSO 进行变量筛选后的边际强度效应曲线。从附图 G4 中可以发现,进行变量筛选后的曲线和基准结果略有不同,尤其是在 43%~46% 的录取率下扩招政策的边际影响出现小幅上升。不过整体而言,扩招政策的边际强度效应依然随着录取率

⁴ Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects" *Journal of Economic Perspectives*, 2014, 28 (2): 29-50.

⁵ 其中惩罚项 $\lambda = 0.04$ 。

的上升而不断下降,并且在55.5%的录取率附近扩招政策的边际影响接近零。因此附图4依然支持前文所述结论。

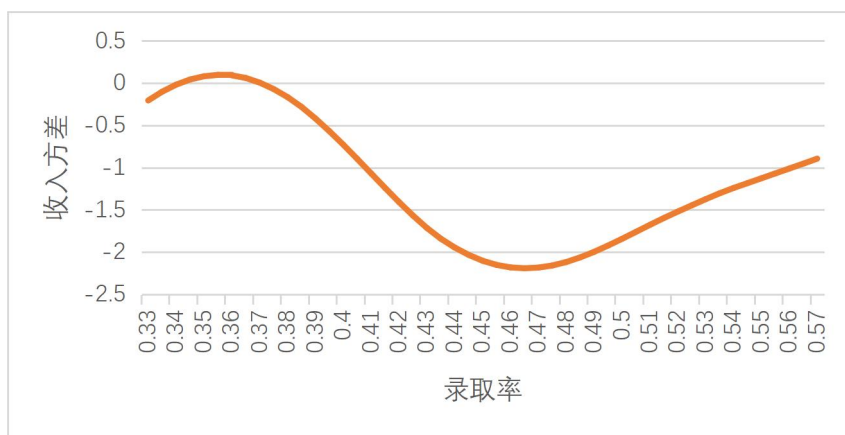


附图G4 变量筛选后的边际强度效应曲线

综合上述,尽管在不同的设定下扩招政策的边际强度效应的大小有所不同,但其随着录取率的上升而下降的趋势均得到保持。因此由于个体异质性的存在,高等教育回报率因人而异,而扩招政策的边际影响随着录取率的上升而下降的结果是稳健的。

H. 扩招政策对收入差距的影响

定理2给出扩招政策对收入方差的识别方式。而收入的方差作为收入差距的一种度量方式,本文通过定理2以及相应的估计方法可以估计出不同录取率下扩招政策对收入差距的影响。附图H1给出扩招政策对收入方差的影响。



附图H1 扩招政策对收入方差的影响

与扩招政策的边际强度效应不同,扩招政策对收入的方法主要是负面影响,即扩招政策缩小了收入差距。我们可以从附表D1中看到产生这一现象的端倪。由附表D1可知,高等教育可以缩小由性别、户籍状况、工作部门、身体健康状况以及资产差异等因素所导致的收入差异,因此高等教育可以通过以上途径降低收入差距,即对收入的方差具有负向影响。不过附图H1表明,当录取率提升到46%左右之后,随着录取率的提升扩招政策降低收入差距的作用逐渐衰弱。

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。