

异质性、选择偏差与教育回报：基于中国微观数据的实证研究*

李雪松 詹姆斯·海克曼**

摘 要

本文根据 2000 年中国的微观数据，运用现代微观计量经济学的分析方法，在考虑异质性和选择偏差的基础上，估计了 20 世纪末中国的教育回报。研究表明：与受教育水平相关的收益在人们中间存在显著的异质性；在当今中国的劳动力市场上存在一种重要的实证现象，即人们根据比较优势原理对教育水平进行选择。传统的普通最小二乘法以及工具变量法都难以对这种选择做出合理的估计，我们的分析框架弥补了上述两种方法的缺陷。2000 年中国 6 个省区城镇青年大学教育的平均回报率为 43% (年均近 11%)。中国在经历二十多年的市场经济改革后，较之 80 年代及 90 年代初期，教育的平均回报有了显著提高，中国的教育和劳动力市场已经开始发挥重要作用。

一、引言

微观数据有两个基本特征：异质性以及缺少与实际相反的状况。由于存在未被观测到的异质性，即使在所有可以被观测到的方面都相同的人们仍然会做出不同的决策、获得不同的收入、选择不同的投资组合。缺少与实际相反的状况引发了数据缺失问题，如果某人实际选择了一种状况，我们就不可能观测到他(她)做出其他选择时的结果。如果我们观测到某大学毕业生的当前收入水平，我们就不可能同时观测到假使他(她)高中毕业就参加工作的当前收入水平。

过去，解决选择和数据缺失问题的方法大多没有考虑异质性，他们一般都假设不同的个人具有同质性。本文的研究充分考虑了个人之间显著的异质性。我们根据中国城镇居民家庭收入与支出调查(CUHIES2000)所公布的横截面微观数据，估计了中国高等教育的异质性回报。对学历做出不同选择的人们有着各自不同的教育回报率。本文的工作建立在由 Heckman 和 Vytlačil (1999, 2000, 2001), Carneiro, Heckman 和 Vytlačil (2001) 以及 Carneiro (2002) 等研究所形成的半参数分析框架基础之上。边际政策效应(MTE)是这些文献的核心概念，这个概念是由 Bjorklund 和 Moffitt (1987) 首先提出的。边际政策效应是指，处于接受或不接受教育临界状态的人最终选择接受教育时的平均回报。这些人根据未被观测到的自身特征选择接受教育的不同水平。

我们的研究表明：对于 2000 年中国 6 省区(广东、辽宁、四川、陕西、浙江和北京)的城镇居民来说，大学 4 年教育的平均政策效应(ATE) — 随机挑选任何一个人接受大学教育后其

* 本文发表在《经济研究》2004 年第 4 期。

** 作者单位：李雪松，中国社会科学院数量经济与技术经济研究所；詹姆斯·赫克曼(James J. Heckman)，美国芝加哥大学经济系。

一生收入将增长 43%(即每年大学教育带来近 11%的回报率)。而用另外两种方法—普通最小二乘法(OLS)和工具变量法(IV)—对 ATE 做出的估计结果分别为 29%和 56% (即每年大学教育分别带来约 7%和 14%的回报率)。总体中存在着显著的回报异质性,估算出的选择偏差达到了-22%。我们的发现与 Carneiro, Heckman 和 Vytlacil (2001)相似:人们根据比较优势原理对教育水平进行选择。最小二乘法(OLS)对真实的平均政策效应(ATE)给出了一个下偏的估计,而工具变量法(IV)则给出了一个上偏的估计。

在经历了二十多年以市场为导向的经济改革后,中国的平均教育回报较之 80 年代和 90 年代初期有了较大幅度的提高。中国的教育和劳动力市场已经开始发挥作用,技能获得的回报也比以前更多了。

本文的余下内容是这样安排的:第二部分描述考虑和不考虑异质性时的收入模型;第三部分对选择偏差、边际政策效应(MTE)进行定义,并给出估计 MTE 的理论框架;第四部分是中国数据的实证分析结果;第五部分进行简要地总结。

二、考虑和不考虑异质性时的模型

考虑如下明瑟尔(Mincer)方程,其参数为常数,这是假设不存在异质性时的传统教育回报模型。

$$\ln Y_i = \beta S_i + \gamma X_i + U_i \quad (1)$$

其中 i 表示不同的个人($i = 1, 2, \dots, n$), $\ln Y_i$ 为收入的对数形式, S_i 表示受教育水平, X_i 为解释变量向量,比如工作年限、工作年限的平方、以及性别、地区、产业、企业所有制等虚拟变量, U_i 是期望为零的随机误差项, β 为教育回报率, γ 为系数向量。

直接运用普通最小二乘法(OLS)对方程(1)进行估计存在一个问题:该模型可能遗漏了个人能力变量 A_i , 它被包含在误差项 U_i 之中。许多实证分析认为 $Cov(A_i, S_i) \neq 0$, 因此 $E(U_i | S_i) \neq 0$ 。此时最小二乘法只能得出有偏、非一致的估计量。Griliches(1977)阐述了对该问题的经典看法。

可是,在现实生活中,大多数数据并不包括对个人能力的度量。于是不少经济学家想方设法运用一些变通的方法,试图来消除或减弱个人能力偏差。许多学者采用工具变量(IV)法,找出一个与 S_i 高度相关但与 U_i 不相关的工具变量 I_i ; 另外一些学者采用固定效应方法,假设双胞胎、兄弟姐妹具有相近的个人能力,对他们进行成对比较;还有一些学者使用个人能力的代理变量,并把这些代理变量作为解释变量包含在 X_i 中。

由于大多数数据并不具备采用固定效应方法所需的足够信息,而且使用该方法还需满足误差项可分离的条件,这种比较可能会使误差更趋严重。找到满意的工具变量也很困难,事实上,大

量有关教育回报和收益问题的文献中普遍使用的工具变量都是无效的，它们往往与被遗漏的个人能力相关(Carneiro 与 Heckman(2002)以及 Carneiro(2002))。

许多实证分析表明：家庭背景往往与个人能力关系密切，良好的家庭氛围通常有助于提高个人能力(Carneiro 和 Heckman, 2003)。本文在实证分析过程中使用父母收入作为个人能力的代理变量。

设想一种比模型 (1) 更为普遍的形式，它考虑了教育的异质性回报，用随机系数的形式表示如下：

$$\ln Y_i = \beta_i S_i + \gamma X_i + U_i \quad (2)$$

其中 β_i 表示存在异质性时的教育回报率，因人而异。 X_i 是由其他解释变量组成的向量，该模型在一种更为普遍的条件下对个人能力偏差进行了校正。

本文着重讨论两种教育水平的选择：高中和大学。令 $S_i = 1$ 表示大学学历， $S_i = 0$ 表示高中学历（没有接受大学教育）。显然存在多种学历选择，我们的分析则是对实际情况的一种简化，这种简化普遍存在于文献中。大量证据表明，在多数情况下，教育回报与受教育年限并不成线性关系，使用传统的收入对数对于学历的回归系数将导致对教育回报率估计的严重偏差。(Heckman, Lochner 和 Todd, 2003)。

两种潜在的选择结果 $(\ln Y_{0i}, \ln Y_{1i})$ 可以表示为：

$$\begin{cases} \ln Y_{0i} = \gamma_0 X_i + U_{0i} & (\text{当 } S_i = 0 \text{ 时}) \\ \ln Y_{1i} = \gamma_1 X_i + U_{1i} & (\text{当 } S_i = 1 \text{ 时}) \end{cases} \quad (3a)$$

$$\quad (3b)$$

对总体而言， $E(U_{0i}|X_i) = 0$ ， $E(U_{1i}|X_i) = 0$ 。

由于存在数据缺失问题，在横截面上想同时获知一个人的 $\ln Y_{0i}$ 和 $\ln Y_{1i}$ 通常是不可能的。我们仅仅可以确定分布 $F(\ln Y_{0i}|X_i, S_i = 0)$ 和 $F(\ln Y_{1i}|X_i, S_i = 1)$ ，而不可能确定分布 $F(\ln Y_{0i}|X_i)$ 和 $F(\ln Y_{1i}|X_i)$ 。由于异质性及选择问题的普遍存在，我们不可能再使用诸如 OLS 和 IV 的传统方法来估计参数。

对结果进行整理，有：

$$\ln Y_i = S_i \ln Y_{1i} + (1 - S_i) \ln Y_{0i} \quad (4a)$$

$$= [(\gamma_1 - \gamma_0) X_i] S_i + \gamma_0 X_i + [U_{0i} + (U_{1i} - U_{0i}) S_i] \quad (4b)$$

$$= [(\gamma_1 - \gamma_0) X_i + (U_{1i} - U_{0i})] S_i + \gamma_0 X_i + U_{0i} \quad (4c)$$

$$= \beta_i S_i + \gamma_0 X_i + U_{0i} \quad (4d)$$

其中

$$\beta_i = (\gamma_1 - \gamma_0) X_i + (U_{1i} - U_{0i}) \quad (5)$$

表示个体 i 的异质性教育回报。当 $\gamma_1 \neq \gamma_0$ (即存在观测到的异质性 $(\gamma_1 - \gamma_0)X_i$), 或 $U_{1i} \neq U_{0i}$ (即存在未观测到的异质性 $(U_{1i} - U_{0i})$) 时, β_i 在总体中是一个变量, 教育回报是一个服从于一定分布的随机变量。在给定 X 的条件下, β_i 的平均值为:

$$\bar{\beta} = E(\beta_i | X_i) = E[(\gamma_1 - \gamma_0)X_i] \quad (6)$$

假设人们根据下述选择规则来决定是否进入大学学习:

$$S_i^* = P_i(Z_i) - U_{si}$$

$$S_i = \begin{cases} 1 & \text{当 } S_i^* \geq 0 \text{ 时} \\ 0 & \text{当 } S_i^* < 0 \text{ 时} \end{cases} \quad (7)$$

其中 S_i^* 为隐藏变量, 代表入学的净收益, Z_i 是可观测到的变量向量 (Z_i 可能包含部分 X_i)。

$P_i = P_i(Z_i)$ 表示参与或接受政策 (比如进入大学学习) 的概率, 它可用概率模型或逻辑模型估计出来。 U_{si} 表示个体 i 在政策选择过程中未被观测到的异质性。

不失一般性, 假设 U_{si} 在 $[0, 1]$ 区间上服从均匀分布 (Heckman 和 Vytlacil, 1999), 对于个体 i 来说, 是否进入大学学习完全取决于观测到的异质性 $P_i(Z_i)$ 与未被观测到的异质性 U_{si} 之间的比较。 U_{si} 越小, 则进入大学学习的可能性越大。

三、选择偏差与边际政策效应

令 $\Delta_i = \ln Y_{1i} - \ln Y_{0i}$ 表示某项政策使个体 i 从 $S_i = 0$ 转变到 $S_i = 1$ 时的经济收益。根据方程 (3a)、(3b) 及 (5), $\Delta_i = \beta_i$ 。根据方程 (3a)、(3b) 及 (6), 普通最小二乘法 (OLS) 估计量的概率极限为:

$$\begin{aligned} p \lim(\hat{\beta}_{OLS}) &= E(\ln Y_i | X_i, S_i = 1) - E(\ln Y_i | X_i, S_i = 0) \\ &= E(\gamma_1 X_i + U_{1i} | X_i, S_i = 1) - E(\gamma_0 X_i + U_{0i} | X_i, S_i = 0) \\ &= \bar{\beta} + [E(U_{1i} | S_i = 1) - E(U_{0i} | S_i = 0)] \end{aligned} \quad (8)$$

(ATE) (偏差)

其中, 平均政策效应 ATE 表示随机挑选一个具有特征 X 的个体所具有的教育回报的均值, 定义如下:

$$ATE = E(\Delta_i | X_i) = E(\beta_i | X_i) = \bar{\beta} \quad (9)$$

如果人们很清楚各自的 U_{si} 且依据 (U_{0i}, U_{1i}) 采取行动, 那么 S_i 将与 U_{0i} 及 U_{1i} 相关, 方程(8)的第二项将不为零。因此普通最小二乘法 OLS 是对 ATE 的有偏估计量。注意到方程(8)也可表示为:

$$\begin{aligned} p \lim(\hat{\beta}_{OLS}) &= E(\ln Y_i | X_i, S_i = 1) - E(\ln Y_i | X_i, S_i = 0) \\ &= E(\beta_i | X_i, S_i = 1) + [E(U_{0i} | S_i = 1) - E(U_{0i} | S_i = 0)] \end{aligned} \quad (10)$$

(TT) (选择偏差)

其中参与者的平均政策效应 TT 表示那些政策参与者（比如进入大学学习）与假设他们没有参与该政策（比如没有接受大学教育）时相比所获得的效应的均值, 定义为:

$$\begin{aligned} TT &= E(\Delta_i | X_i, S_i = 1) \\ &= E(\beta_i | X_i, S_i = 1) \\ &= \bar{\beta} + E(U_{1i} - U_{0i} | S_i = 1) \\ &= ATE + E(U_{1i} - U_{0i} | S_i = 1) \end{aligned} \quad (11)$$

(分类效应)

选择偏差 $E(U_{0i} | S_i = 1) - E(U_{0i} | S_i = 0)$ 表示事实上的政策参与者（比如进入大学学习）假设他们未参加项目时的不可观测变量与真正未参加项目者（比如未接受大学教育）的不可观测变量之差的均值, 也即大学毕业生如果他们高中毕业就参加工作的当前收入与真正高中毕业生的当前收入之间的不可观测变量差异的均值。

分类效应 $E(U_{1i} - U_{0i} | S_i = 1)$ 表示那些政策参与者（比如进入大学学习）基于他们自身不可观测变量所带来的平均效应。分类效应与选择偏差之和构成了(8)式中的偏差。

在存在异质性和选择偏差的情况下, 传统的 IV 估计方法通常不能正确估计这些参数。找出与 S_i 相关但与 U_{0i} 甚至 $U_{1i} - U_{0i}$ 不相关的工具 I_i 并不足以确定 $\bar{\beta}$, 因为:

$$\begin{aligned} p \lim \hat{\beta}_{IV} &= \frac{Cov(I_i, \ln Y_i)}{Cov(I_i, S_i)} = \bar{\beta} + \frac{Cov(I_i, U_{0i})}{Cov(I_i, S_i)} + \frac{Cov[I_i, (U_{1i} - U_{0i})S_i]}{Cov(I_i, S_i)} \\ &= \bar{\beta} + \frac{Cov[I_i, (U_{1i} - U_{0i})S_i]}{Cov(I_i, S_i)} = \bar{\beta} + \frac{Cov[I_i, (U_{1i} - U_{0i}) | S_i = 1]P_i}{Cov(I_i, S_i)} \end{aligned} \quad (12)$$

其中 $P_i = \Pr(S_i = 1)$ 是政策的参与率。在异质性和选择偏差都存在时 $U_{1i} \neq U_{0i}$, $U_{1i} - U_{0i}$ 依赖于 S_i , 方程(12)中的第二项不为零, 所以 $p \lim(\hat{\beta}_{IV}) \neq \bar{\beta}$ 以致 IV 不是一致估计量。只有在 $U_{1i} - U_{0i} = 0$ (即未观测到的异质性和选择偏差都不存在) 或 $U_{1i} \neq U_{0i}$ 但 $U_{1i} - U_{0i}$ 独立于 S_i (即存在不可观测到的异质性, 但不存在选择偏差) 等一些非常特殊的情况下, 方程(12)的第二项才为零。此时, IV 才可能成为 $\bar{\beta}$ 的一致估计量 (Heckman, 1997; Heckman、Navarro-Lozano, 2003)。

因此, 当存在异质性和选择问题时, 普通最小二乘法 OLS 和工具变量法 IV 都难以对教育回

报率给出一致的估计。可是，在一定的假设条件下（Heckman 和 Vytlacil (1999, 2000, 2001), Carneiro, Heckman 和 Vytlacil (2001), Carneiro (2002) 和 Navarro-Lozano (2002)），仍可能采用局部工具变量 (*LIV*) 法，根据边际政策效应 (*MTE*) 来确定教育的异质性回报。其中 *MTE* 是在 X_i 及未观测到的异质性 U_{si} 给定的情况下对 $\ln Y_{1i}$ (与 $\ln Y_{0i}$ 相比) 的平均支付意愿 (*WTP*)，

定义如下：

$$\begin{aligned} MTE(X_i = x, U_{si} = u_s) &= E(\Delta_i | X_i = x, U_{si} = u_s) = E(\beta_i | X_i = x, U_{si} = u_s) \\ &= (\gamma_1 - \gamma_0)x + E(U_{1i} - U_{0i} | U_{si} = u_s). \end{aligned} \quad (13)$$

MTE 运用下式进行估计，其中局部工具变量 *LIV* 可以通过推广了的半参数方法进行计算 (Heckman, 2001)：

$$MTE(X_i = x, U_{si} = P_i = p) = LIV(X_i = x, P_i = p) = \frac{\partial E(\ln Y_i | X_i = x, P_i = p)}{\partial p} \quad (14)$$

为了表达的方便，在下面公式中我们隐去以 X 为条件的表述，Heckman、Vytlacil (1999, 2000, 2001)、Carneiro (2002) 建立了一套通过 *MTE* 来计算其他政策效应的方法：

$$ATE = \int_0^1 MTE(u_s) du_s \quad (\text{平均政策效应})$$

$$TT = \int_0^1 MTE(u_s) h_{TT}(u_s) du_s \quad (\text{参与者的平均政策效应})$$

$$TUT = \int_0^1 MTE(u_s) h_{TUT}(u_s) du_s \quad (\text{未参与者的平均政策效应})$$

其中各个权数如下：

$$h_{ATE}(u_s) = 1$$

$$h_{TT}(u_s) = \frac{1 - F_P(u_s)}{E(P_i)} = \frac{\int_{u_s}^1 f(p) dp}{E(P_i)}$$

$$h_{TUT}(u_s) = \frac{F_P(u_s)}{E(1 - P_i)} = \frac{\int_0^{u_s} f(p) dp}{E(1 - P_i)}.$$

未参与者的平均政策效应 (*TUT*) 是指没有参与该政策的人们（比如没有进入大学学习）与假如他们参与（比如进入大学）时相比所产生的政策效应的均值，定义如下：

$$\begin{aligned} TUT &= E(\Delta_i | X_i, S_i = 0) = E(\beta_i | X_i, S_i = 0) \\ &= \bar{\beta} + E(U_{1i} - U_{0i} | S_i = 0). \end{aligned} \quad (15)$$

四、中国数据的实证分析结果

我们使用的 2000 年中国城镇居民家庭收支调查 (CUHIES) 数据来源于国家统计局城镇经济社会调查总队。这项调查包含 1992—2003 年一系列横截面数据，目前还在继续，城镇数据由城镇居民总体随机抽取构成。

我们有 2000 年中国 6 个省区城镇地区的数据，这 6 个省区为：广东、辽宁、四川、陕西、浙江和北京。2000 年北京、广东、浙江三个省区的居民平均收入高于全国的平均水平，四川、辽宁、陕西三省低于全国的平均水平。6 个省区城镇居民平均收入为 7627 元，高于全国平均水平的 6280 元。

6 个省区样本规模为 4250 个家庭。数据提供户主、配偶、子女、父母等每一位家庭成员丰富的信息，包括年龄、性别、受教育程度、产业、企业所有制情况、职业、工龄、年收入等。样本中的受教育程度有 7 种选择：大学、大专、职业培训学校、高中、初中、小学及其他。

为便于研究，我们把 6 个省区每个家庭具有大学或大专学历以及高中学历的子女列出。他们在 2000 年必须有工作，而且取得一定的收入。这样处理后样本由 587 个群体（331 名男性，256 名女性）组成，其中 273 人拥有大学本科或大专文凭，其余 314 人为高中学历。样本的平均年龄为 26.3 岁，主要是青年人，因此，我们的分析是中国新锐一代在其生命周期早期的收入水平和教育回报。

表 1. 估计出的明瑟尔模型

变量	普通最新二乘法 OLS		工具变量法 IV [#]	
	系数	标准差	系数	标准差
截距	8.3189	0.1493	8.3040	0.1552
大学 4 年教育 e	0.2929	0.0630	0.5609	0.1695
工龄	0.0380	0.0194	0.0196	0.0202
工龄平方	-0.0016	0.0010	-0.0007	0.0010
以千元为单位的父母收入	0.0117	0.0020	0.0098	0.0023
男性	0.1537	0.0602	0.1439	0.0607
居住于广东省	0.7543	0.1255	0.7908	0.1267
居住于辽宁省	0.2693	0.1085	0.3142	0.1092
居住于四川省	0.2278	0.1181	0.2759	0.1192
居住于北京市	0.7246	0.1241	0.7775	0.1256
居住于浙江省	0.6241	0.1297	0.6739	0.1314
在国有企业工作	-0.3679	0.0855	-0.3873	0.0868
在集体企业工作	-0.4786	0.1288	-0.5890	0.1298
在私营企业工作	-0.4649	0.1179	-0.5304	0.1179
在 IND_CON* 产业工作	-0.2793	0.0788	-0.3048	0.0792
在 TRA_COM* 产业工作	-0.4512	0.1762	-0.4645	0.1779
在 SPO_SOC* 产业工作	-0.2880	0.0900	-0.3106	0.0905
在 FIN_INS* 产业工作	-0.3220	0.1050	-0.3327	0.1061

注*：IND_CON 表示工业、地质勘测和人口普查、建筑产业；TRA_COM 表示交通、运输、邮政电信、商业、家政、材料供应产业；SPO_SOC 表示卫生、体育和社会福利产业；FIN_INS 表示金融和保险产业。

注*：使用大学教育的概率作为接受四年大学教育的工具变量。

普通最小二乘法 *OLS* 和工具变量法 *IV* 的估计结果见表 1。我们使用进入大学的概率作为工具变量。两种方法估计出的教育回报分别为 29%和 56%，年均教育回报率分别为 7.25%和 14%。OLS 估计量比起由邹至庄 (CHOW, 2001) 报告的早些时期 (20 世纪 80 年代和 90 年代早期) OLS 估计值要大。收入方程中的变量包括：工作年限、工作年限的平方、个人能力代理变量 (在本文中，我们使用父母收入作为代理变量)，以及诸如性别、居住省区、企业所有制形式、产业等虚拟变量。

用逻辑模型对大学入学概率进行估计的结果见表 2，逻辑模型的一般形式及根据它导出的边际效应定义如下：

$$\Pr(S = 1 | X) = \frac{\exp(\beta'X)}{1 + \exp(\beta'X)} = \Lambda(\beta'X) \quad (16)$$

$$\text{边际效应} = \frac{\partial \Pr(S = 1 | X)}{\partial X} = \Lambda(\beta'X)[1 - \Lambda(\beta'X)]\beta$$

表 2. 接受高等教育的逻辑模型

变量	系数	标准差	平均边际效应
截距	-4.7370	0.7305	-
父亲受教育年限	0.1017	0.0297	0.0211
母亲受教育年限	0.0605	0.0342	0.0126
以千元为单位的父母收入	0.0190	0.0069	0.0040
生于 1964 年之前	2.0008	0.7969	0.4159
生于 1964 年	1.7285	0.9189	0.3593
生于 1965 年	3.3423	0.8257	0.6947
生于 1966 年	3.1813	0.8552	0.6613
生于 1967 年	1.8455	1.1126	0.3836
生于 1968 年	2.9030	0.8161	0.6034
生于 1969 年	2.2569	0.7941	0.4691
生于 1970 年	1.5076	0.7534	0.3134
生于 1971 年	3.0771	0.7138	0.6396
生于 1972 年	2.6424	0.7183	0.5492
生于 1973 年	2.5395	0.6809	0.5279
生于 1974 年	2.7740	0.6753	0.5766
生于 1975 年	2.7931	0.6763	0.5806
生于 1976 年	2.8634	0.6669	0.5952
生于 1977 年	2.5890	0.6672	0.5381
生于 1978 年	2.5572	0.6656	0.5315
生于 1979 年	1.3631	0.7636	0.2833

在逻辑模型中，我们使用父母双方的教育水平、父母收入以及个人出生年份作为大学入学概率的解释变量。表 2 的最后一列是各个解释变量的平均边际效应。图 1 给出了估计出的大学入学概率的密度函数 $\Pr(S=1)$ 。

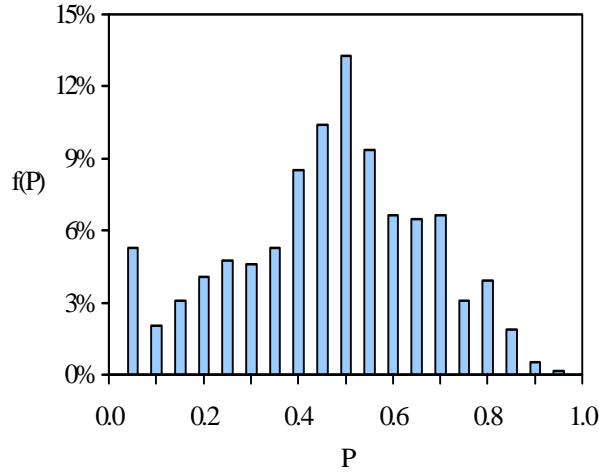


图 1. $P(S=1)$ 的密度分布

表 3 和图 2 是采用半参数方法对边际政策效应 MTE 的估计结果。我们在收入方程中使用父母收入来对个人能力进行代理。表 3 包含了由局部线性回归估计出的方程 (3a) 和 (3b) 中的系数。图 2 是估计出的边际政策效应 MTE，它是未观测到的异质性 u_s 的函数。MTE 随 u_s 增大而减小。这表明： u_s 值较小的个人（即根据决策规则 (7) 更有可能进入大学）拥有较高的教育边际回报。 u_s 值最大、最没有可能进入大学的个人拥有最低的教育边际回报。图 2 表明中国教育回报存在着显著的异质性。递减的 MTE 曲线还表明，匹配法、传统的 OLS、IV 方法都难以正确估计有关的政策效应。

表 4 给出了不同政策效应参数的对比。ATE 为 43%（每年近 11%），表示随机抽取的个人四年大学教育的平均回报。OLS 估计值为 29%（每年约 7%），是对 ATE 的下偏估计。中国的实证分析结果表明： $IV > ATE > OLS$ 。选择偏差在估计中国教育回报过程中是非常重要的，达到了 -22%。TT，参与者的政策效应与 TUT，未参与者的政策效应分别为 51% 和 36%。分类收益为 8%，是一个较大的正数，表明了比较优势的存在。分类收益还可表示为：

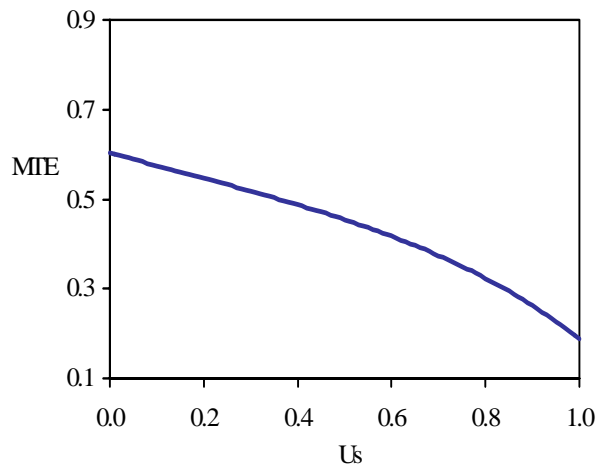
$$\text{分类收益} = E(\beta_i - \bar{\beta} | X_i, S_i = 1) = E(\beta_i | X_i, S_i = 1) - \bar{\beta} = TT - ATE \quad (17)$$

表 3. 局部线形回归估计结果

(高斯核估计, 带宽=0.4)

变量	高中		大学	
	γ_0	标准差	γ_1	标准差
工龄	0.0360	0.0225	0.0141	0.0278
工龄平方	-0.0013	0.0011	-0.0009	0.0013
以千元为单位的父母收入	0.0188	0.0038	0.0077	0.0038
男性	0.1365	0.0723	0.1913	0.0777
居住于广东省	0.5712	0.1961	0.8853	0.1590
居住于辽宁省	0.1901	0.1263	0.3929	0.1049
居住于四川省	0.2612	0.1364	0.2296	0.1081
居住于北京市	0.7122	0.1695	0.7971	0.1301
居住于浙江省	0.6930	0.1551	0.5461	0.1744
在国有企业工作	-0.3368	0.1188	-0.4471	0.1093
在集体企业工作	-0.6060	0.2065	-0.5868	0.1771
在私有企业工作	-0.4205	0.1511	-0.6256	0.1677
在IND_CON*产业工作	-0.2297	0.0821	-0.3978	0.0990
在TRA_COM*产业工作	-0.3527	0.1318	-0.5040	0.1557
在SPO_SOC*产业工作	-0.3702	0.1282	-0.3040	0.1202
在FIN_INS*产业工作	-0.3345	0.1560	-0.3543	0.1331

注*:IND_CON表示工业、地质勘测和人口普查、建筑产业; TRA_COM表示交通、运输、邮政电信、商业、家政、材料供应产业; SPO_SOC表示卫生、体育和社会福利产业; FIN_INS表示金融和保险产业。



方程中包含父母收入作为个人能力代理变量, 带宽=0.4

图 2. 边际政策效应

表 4. 不同参数比较

参数	估计值
OLS	0.2929
IV ¹	0.5609
ATE	0.4336
TT	0.5149
TUT	0.3630
偏差 ²	-0.1407
选择偏差 ³	-0.2220
分类收益 ⁴	0.0813

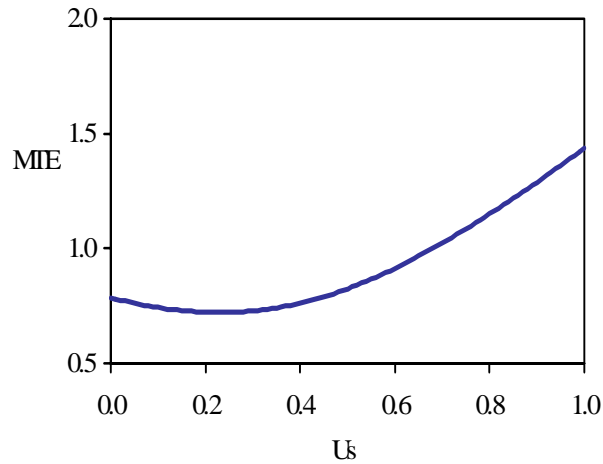
注：¹使用大学入学概率作为工具变量

²偏差 = $OLS - ATE$

³选择偏差 = $OLS - TT$

⁴分类收益 = $TT - ATE$

表 4 和图 2 都反映出，2000 年中国 6 省区城镇青年大学教育平均回报相当可观，但个体间存在着显著的异质性、存在着正的分类收益以及负的选择偏差。



方程中不含父母收入变量，带宽=0.33

图 3. 边际政策效应

为了检验将个人能力代理变量引入收入方程的重要性，我们在去除收入方程中父母收入一项后，重新估计了政策效应。结果在图 3 中。 MTE 随 u_s 递增，且平均值也明显高于图 2。由此可

见，忽略个人能力（或其代理变量）将导致对边际政策效应以及估计出的教育回报的上偏。如果不引入个人能力（父母收入）代理变量，教育回报估计值将上升到令人难以置信的较高水平，同时也改变了估计出的 *MTE* 的曲线形状。这些与 Carneiro (2002) 运用美国数据得到的结论基本一致。

五、结束语

本文使用微观数据来估计中国的高等教育回报。我们揭示了考虑异质性和选择偏差的重要性。忽略这两点将导致有偏、不一致的估计值（比如在使用 OLS、IV 参数法时就出现了此种情况）。我们还揭示出在收入方程中使用个人能力代理来识别教育回报的重要性。忽略这一点将导致不真实的对教育回报的过高估计。

2000 年，对于中国 6 省区的城镇青年来说，大学四年教育平均回报为 43%（即每年近 11%）。对于大学生来说，该回报率更高。这些结果高于明瑟尔模型中传统的 OLS 估计值，同时也高于早些时候其他一些文献的 OLS 估计值。这些结果表明，在经历了二十多年的以市场为导向的经济改革后，中国的平均教育回报较之 20 世纪 80 年代和 90 年代早期有了显著提高。

主要参考文献：

- Bjorklund, A. and R. Moffitt (1987), “The Estimation of Wage Gains and Welfare Gains in Self-Selection Models”, *Review of Economics and Statistics*, 69:42-49.
- Carneiro, P. (2002), “Heterogeneity and Selection in the Return to Schooling: Implications for Education Policy Evaluation”, Ph.D. Dissertation, University of Chicago.
- Carneiro, P. and J. Heckman (2002), “The Evidence on Credit Constraints in Post-Secondary Schooling”, *Economic Journal*, Volume 112, issue 482, pp:705-734.
- _____. (2003), “Human Capital Policy,” in J. Heckman, A. Krueger, *Inequality in America: What Role for Human Capital Policies*, Cambridge, MA: MIT Press.
- Carneiro, P., J. Heckman and E. Vytlačil (2001), “Estimating the Return to Education When It Varies Among Individuals”, Working paper, University of Chicago.
- China's Urban Household Income and Expenditure Survey (CUHIES) (2000), National Bureau of Statistics, Urban Socio-Economic Survey Organization.
- Chow, G. (2001), *China's Economic Transformation*. Malden, MA: Blackwell Publishers.
- Griliches, Z. (1977), “Estimating the Returns to Schooling: Some Econometric Problems,” *Econometrica*, Volume 45, number 1, pp. 1-22.
- Heckman, J. (2001), “Microdata, Heterogeneity and Econometric Policy Evaluation”, Nobel Memorial Lecture in Economic Sciences, *Journal of Political Economy*, 109(4), 673-748.
- Heckman, J., L. Lochner and P. Todd. (2003) “Fifty Years of Mincer Earnings Functions,” Unpublished

manuscript. University of Chicago.

Heckman, J. and S. Navarro-Lozano. (2003), "Using Matching, Instrumental variables and Control Functions to Estimate Economic Choice Models" forthcoming in *Review of Economics and Statistics*. October.

Heckman, J. and E. Vytlacil (1999), "Local Instrumental variable and Latent variable Models for Identifying and Bounding Treatment Effects", *Proceedings of the National Academy of Sciences*, 96:4730-4734.36.

_____. (2000), "Local Instrumental variables", in C. Hsiao, K. Morimune, and J. Powells, (eds.), *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya*, Cambridge: Cambridge University Press, 1-46.

_____. (2001), "Structural Equations, Treatment Effects and Econometric Policy Evaluation", Fisher Schultz Lecture World Congress of the Econometric Society.

National Bureau of Statistics (NBS). (2001), *China Labor Statistical Yearbook*. Beijing: China Statistics Press.