

理解 R&D 投入回报之谜： 引入不确定性、非线性和企业异质性

尹 恒¹, 柳 荻²

(1. 中国人民大学 国家发展与战略研究院, 北京 100872;

2. 北京师范大学 经济与工商管理学院, 北京 100875)

摘 要:既有研究发现企业 R&D 投入的回报不高甚至为负, 这与 R&D 投入对经济增长具有重要作用的直觉背道而驰。其原因在于传统以知识资本模型为基础的研究, 其估计的是平均 R&D 投入回报, 平滑掉了 R&D 投入回报的企业异质性。文章建立了结构估计模型正式处理了 R&D 投入回报的不确定性、非线性以及企业异质性。与既有结构模型的区别在于, 文章在模型中考虑了溢价的企业异质性, 并明确处理了影响企业需求的因素。研究结果表明: R&D 投入的回报确实是不确定的和非线性的, 从而具有十分明显的企业异质性; 个别企业的 R&D 投入回报甚至可以高出平均回报数倍, 这足以为企业提供很强的 R&D 投入激励。R&D 投入回报的企业异质性理解平均意义上 R&D 投入回报较低与宏观层面 R&D 投入对经济增长具有重要作用的现实并存的钥匙。因此, 我们有充分的理由相信, R&D 投入及由此带来的生产率改善, 能够成为新常态下推动中国经济持续发展的主要力量。

关键词: R&D 投入回报; 企业生产率; 企业异质性; 结构估计

中图分类号: F424.3 **文献标识码:** A **文章编号:** 1001-9952(2016)09-0121-12

DOI: 10.16538/j.cnki.jfe.2016.09.011

一、引言和文献回顾

企业 R&D 投入的回报有多大? 张海洋(2005)运用中国 1999—2002 年 34 个制造业行业的面板数据研究发现, R&D 投入对生产率的作用不显著, 甚至为负, 进而认为 R&D 投入对企业生产率的提高并没有明显作用。Hu 等(2005)运用中国 1995—1999 年每年约 10 000 个大中型制造企业数据研究发现, R&D 投入产出弹性为 0.027—0.029, 进一步分解为高科技企业和非高科技企业, 发现后者的 R&D 投入对生产率并没有显著影响。Jefferson 等(2006)利用中国 1997—1999 年 5 451 个大中型制造企业面板数据研究发现, R&D 投入产出弹性约为 0.24; 然而使用工具变量法时 R&D 投入的产出弹性却变为 -0.045。李小平和朱钟棣(2006)运用 1998—2003 年 32 个制造业行业的面板数据研究发现, R&D 投入对行业技术进步、技术效率和全要素生产率的增长起阻碍作用。吴延兵(2006)运用中国 2002 年 537 个四位数制造业行业数据进行研究, 得到了 R&D 投入对生产率有显著的促进作用的结论。不过在含有中间投入的生产函数中, R&D 投入的产出弹性只有 0.01—0.02。周亚虹等(2012)根据全部国有和规模以上非国有工业企业数据库中 2005—2007 年的企业累积

收稿日期: 2015-12-10

基金项目: 国家自然科学基金(71173019, 71373026)

作者简介: 尹 恒(1969—), 男, 湖南华容人, 中国人民大学国家发展与战略研究院教授, 博士生导师;

柳 荻(1987—), 女, 黑龙江依安人, 北京师范大学经济与工商管理学院博士研究生。

R&D投入量构建了知识资本存量,对2007年的生产函数进行简单OLS回归,所得到的R&D投入的产出弹性为0.02,使用工具变量的回归结果为0.04—0.06。对其他国家R&D投入回报的估计也得到类似的结论。例如,Griliches(1979)最早估计了美国制造业的R&D投入产出弹性在0.07左右;Griffith和Reenen(2006)对英国的估计结果在0.03左右。偏低甚至为负的R&D投入回报,似乎指向一个结论:R&D投入对企业发展来说无关紧要。然而,经济增长理论却赋予企业R&D投入十分重要的意义,甚至认为它是增长的发动机(Romer,1990)。本文称这个理论与经验证据间的鲜明对照为“R&D投入回报之谜”。

R&D投入回报的不确定性、非线性和企业异质性是理解“R&D投入回报之谜”的关键环节。一方面,即使平均而言R&D投入的回报并不高,但由于存在企业异质性,不排除某些企业在特定情况下R&D投入存在高额的回报,这能够为这些企业进行R&D投入提供充分的激励。另一方面,R&D投入的直接成果,即新工艺和新产品具有很强的非竞争性,对于整个行业甚至经济具有很强的外溢作用。个别企业由高额预期回报驱动的R&D活动,可能获得行业生产技术和产品方面的重大突破,甚至可以成为推动整个经济增长的重要动力(Aghion和Howitt,1992)。因此,只要R&D投入回报具有较强的企业异质性,即使平均而言R&D投入的回报并没有很高的吸引力,但也不妨碍R&D投入在宏观上对于整个经济增长具有重要意义。

然而,评估R&D投入回报的标准模型——Griliches(1979)的知识资本模型,没有考虑不确定性、非线性和企业异质性。Griliches(1979)提出在生产函数中引入技术知识当前状态的测度,它取决于企业当前和过去的R&D投入即知识资本。知识资本作为与资本和劳动对应的另一生产要素,其产出弹性可以度量R&D投入的产出回报。经验研究中一般运用永续盘存法,根据企业各年R&D支出流量序列构建出知识资本存量的代理变量,然后运用双对数模型回归出生产函数中各要素的产出弹性系数。Griliches提供了可操作的工具,随后沿着这一思路考察R&D投入回报的文献大量涌现,并且Mairesse和Sassenou(1991)、Griliches(1995、1998)、Hall等(2010)对此作了比较全面的综述。知识资本模型目前仍然是评估R&D投入产出绩效的主要工具,国内也有一些研究运用这一方法(吴延兵,2006;陆国庆,2011;周亚虹等,2012)。很明显,在知识资本模型中,R&D投入的产出弹性如同生产函数的其他参数,在构建上,行业内所有企业都是相同的。因此,如果两个企业的R&D投入序列相同,其对产出绩效的效应必然完全一样。可见知识资本框架很难令人满意地处理R&D投入回报的不确定性、非线性和企业异质性问题。

评估R&D投入产出绩效的另一常见做法是先估计企业生产率,然后用R&D投入(加上其他控制变量)对其进行回归。^①其中估计生产率的方法除了生产函数回归(Solow 剩余)外,还包括数据包络方法(DEA)、随机前沿方法(SFA)和指数方法。如张海洋(2005)、李小平和朱钟棣(2006)均使用基于DEA的Malmquist生产率测度。这种方法在第一步估计生产率时没有充分考虑随机因素和不确定性,第二步回归在设定上假定R&D投入的产出(生产率)弹性对于所有企业都是同质的,从而仍然没有充分处理R&D投入回报的不确定性、非线性和企业异质性问题。而且,这种两步法需要应对明显的内生性挑战:R&D投入与生产率可能互为因果、可能存在同时影响两者的不可观测因素。宗庆庆等人(2013)、毛德凤等人(2013)尝试运用政策评估中的广义倾向得分匹配方法(PSM)估计R&D投入对生产率

^①容易证明,生产率的R&D投入弹性就是产出的R&D投入弹性,也即本文分析的R&D投入回报。

的处理效应。这仍然需要假定 R&D 投入和生产率间的确定性关系。

本文借鉴 Aw 等(2011)以及 Doraszelski 和 Jaumandreu(2013)的思路,将 R&D 投入引入企业生产率演化的马尔科夫过程,从而使其对生产率的分布产生影响,进而建立结构模型正式考虑 R&D 投入回报的不确定性、非线性和企业异质性问题。在他们模型的基础上,我们进行了改进,明确考虑了溢价的企业异质性,并明确处理了影响企业需求的因素。正如 Klette 和 Griliches(1996)所指出的,不完全竞争环境下企业层面的价格还受到需求方面因素的影响,用行业水平的价格指数平减企业销售收入作为产出的代理指标,会将企业的需求异质性引入生产率估计方程,从而造成生产函数参数与生产率估计的严重偏误。^① 考虑到不完全竞争环境是最近生产率文献的一个热点。例如,De Loecker(2011)在分析比利时纺织品市场贸易保护的下降对企业生产率的影响时,同时考虑了生产率异质性和需求异质性。De Loecker 和 Warzynski(2012)在分析溢价与企业出口行为间的关系时,其使用企业层面数据估计溢价的方法就是建立在不完全竞争企业短期成本最小化的基础之上。我们也将 Olley 和 Pakes(1996)、Levinsohn 和 Petrin(2003)以及 Akerberg 等(2006)的企业生产率结构估计思路扩展到不完全竞争环境下,同时考虑企业需求异质性。

基于 GMM 的检验表明,在企业生产率演化的马尔科夫过程中引入 R&D 投入是很有必要的,企业本期 R&D 投入确实会对其下期生产率产生显著影响。进一步的检验也显示,R&D 投入对企业生产率的这种影响具有不可忽视的不确定性、非线性和企业异质性。虽然平均意义上 R&D 投入的回报与现有文献的估计并没有明显差别,然而其确实显示出高度的不确定性和企业异质性。对所估计的生产率进行进一步分析可以得到十分明确的结果:平均而言,所有行业本期进行研发投入的企业,其下期生产率显著高于其他企业,其生产率的不确定性(方差)也更高;Kolmogorov-Smirnov 检验显示,R&D 投入会显著改变行业内企业的生产率分布,而且对于所有行业,进行 R&D 投入的企业下期生产率分布(随机占优意义上)显著优于没有进行 R&D 投入的企业。这些结果表明,在考虑了影响的不确定性、非线性和企业异质性之后,中国制造业所有行业的 R&D 投入都显著地促进了企业生产率发展。R&D 投入回报的企业异质性可以解释为什么平均意义上 R&D 投入的较低回报与宏观上 R&D 投入具有重要意义能够共存。

文章余下的结构安排是:第二部分提出结构模型和估计思路;第三部分是估计结果;第四部分对全文进行总结。

二、R&D 投入回报的结构估计

企业 i 在 t 期资本存量为 K_{it} ,企业当期投入劳动 L_{it} 和中间材料 M_{it} ,其产出 Y_{it} 为:

$$Y_{it} = K_{it}^{\beta} L_{it}^{\beta} M_{it}^{\beta} e^{\omega_{it}} e^{\mu_{it}} \quad (1)$$

ω_{it} 为企业异质的生产率。同样的资本、劳动和中间材料投入,其产出因 ω_{it} 的差异而不同。 t 期初企业能够观测到自己的生产率水平,但 ω_{it} 对于外部分析人员而言是不可观测的。^② μ_{it} 为企业自己也不可观测的纯粹的产出扰动。

^①Foster 等(2008)也指出,不完全竞争市场上企业异质的需求因素会造成 Olley 和 Pakes(1996)型代理变量方法的遗漏变量偏差。也就是说,代理变量方法要求分析者所使用的不可观测到的企业生产率与其可观测到的作为生产率代理的变量必须一一对应。如果还存在其他企业层面的不可观测因素影响代理变量,Olley 和 Pakes(1996)型代理变量方法就会失效。如果忽视了在不完全竞争市场上的需求因素,显然会造成代理变量方法面临严重的问题。

^②由于缺乏企业层面的价格信息,实际产出也是外部分析人员不可观测的。

自 Olley 和 Pakes(1996)发表了开创性的论文以来,将企业生产率演化设定为一阶马尔科夫过程成为文献中的标准做法(例如 Levinsohn 和 Petrin, 2003; Akerberg 等 2006; Wooldridge, 2009; De Loecker, 2011; De Loecker 和 Warzynski, 2012)。借鉴 Aw 等(2011)和 Doraszelski 和 Jaumandreu(2013),进一步设企业 R&D 投入会影响其生产率演化的马尔科夫过程。具体地:

$$\omega_{it} = E(\omega_{it} | \omega_{it-1}, r_{it-1}) = \gamma_t + g(\omega_{it-1}, r_{it-1}) + \xi_{it} \quad (2)$$

r_{it-1} 为企业上一期的 R&D 投入(对数)。小写字母为大写字母的对数(下同)。除共同的时间趋势 γ_t 外,企业 t 期的实际生产率 ω_{it} 包括两个部分:企业在 $t-1$ 期进行 R&D 投入时的期望生产率 $g(\omega_{it-1}, r_{it-1})$ 和 R&D 投入后生产率的随机扰动 ξ_{it} 。企业 t 期的期望生产率取决于已经取得的生产率水平(ω_{it-1}),也取决于上期 R&D 投入 r_{it-1} 。这里,我们使用研发投入的流量作为研发行为的测度,主要是考虑到在模型的假设中,企业当期的生产率除了在很大程度上延续前一期生产率的惯性外,同时也与前一期的研发行为密切相关。研发投入的存量可能在长期决定企业的平均效率水平,而研发投入的流量则会更加直接地影响到企业效率演化的动态路径。另外,受到我国工业企业数据库的限制,仅凭 2001 年和 2005—2007 年这四年的研发投入流量数据也很难准确估测一个企业的研发投入存量。随机扰动 ξ_{it} 代表企业从 R&D 投入形成实际产出绩效过程中所面临的技术和市场等不确定性。企业 R&D 投入对生产率的影响及其不确定性、非线性和企业异质性正是通过这种方式得以模型化。企业是了解其 $g(\omega_{it-1}, r_{it-1})$ 的,但它对于外部分析人员而言却是不可观测的。

根据式(1)和式(2),可以看到 $\frac{\partial g(\omega_{it-1}, r_{it-1})}{\partial r_{it-1}}$ 即为 R&D 投入的期望产出弹性(边际),本文称为 R&D 投入的边际弹性。另外,如果我们沿用先估计出生生产率 ω_{it} ,然后用 r_{it-1} 对其进行回归的传统思路, r_{it-1} 的系数即为 R&D 投入的实际产出弹性(平均),本文称为 R&D 投入的平均弹性。从而在这个意义上,R&D 投入对企业生产率的影响和 R&D 投入的产出绩效是一回事。估计出了 ω_{it} 和条件期望函数 $g(\omega_{it-1}, r_{it-1})$, R&D 投入的产出绩效评估问题就迎刃而解了。

很多企业(至少在部分年度)不进行 R&D 投入。例如在我们分析的样本中,R&D 投入为正的企业所占比重一般只有 4.85%—22.11%。基于这样的事实,我们借鉴 Doraszelski 和 Jaumandreu(2013)的研究,允许有 R&D 投入和没有 R&D 投入企业的生产率的演化规律存在差异,即有不同的条件期望函数 $g(\omega_{it-1}, R_{it-1})$ 。具体地:

$$g(\omega_{it-1}, r_{it-1}) = 1(R_{it-1} = 0)(g_{00} + g_{01}(\omega_{it-1})) + 1(R_{it-1} > 0)[g_{10} + g_{11}(\omega_{it-1}, r_{it-1})] \quad (3)$$

$1(R_{it-1} = 0)$ 和 $1(R_{it-1} > 0)$ 分别为不进行 R&D 投入和进行 R&D 投入的示性函数。

与 Doraszelski 和 Jaumandreu(2013)的研究不同,我们明确处理企业需求价格弹性的异质性,并控制引起需求曲线移动的因素。具体地,设市场是垄断竞争的,企业 i 面临如下 Dixit-Stiglitz 型需求函数:^①

$$P_{it} = \Phi_t (Y_{it})^{-\frac{1}{\eta_i}} e^{h(A_{it})} e^{\zeta_{it}} \quad (4)$$

P_{it} 为企业产出价格, Φ_t 反映 t 期总的市场特征, $\eta_i > 0$ 为企业异质的需求弹性(绝对

^①这是生产率估计中广泛使用的需求函数,例如 Aw 等(2011)和 DeLoecker(2011)。

值)。企业可以通过加强广告营销、改善售后服务等销售努力(A_{it}),提高消费者对其产品的评价,从而移动其需求曲线。 $h(\cdot)$ 为未知函数。 ζ_{it} 为需求扰动。

这样,观察到的企业实际销售收入(取对数)为:

$$rev_{it} = \varphi_t + \left(1 - \frac{1}{\eta_i}\right) (\beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it}) + h(A_{it}) + \left(1 - \frac{1}{\eta_i}\right) \omega_{it} + v_{it} \quad (5)$$

$v_{it} \equiv \left(1 - \frac{1}{\eta_i}\right) \mu_{it} + \zeta_{it}$ 。企业每期劳动和中间材料可以充分(无成本)调整,从而当期劳动和中间材料选择只影响当期利润。企业 t 期选择 L_{it} 和 M_{it} 最大化当期税后利润,根据中间材料的一阶条件:

$$\omega_{it} = \frac{\eta_i}{\eta_i - 1} \left[\begin{aligned} & p_{M_t} - \varphi_t - \ln \beta_m - \ln \left(1 - \frac{1}{\eta_i}\right) + \left(\frac{1}{\eta_i} - 1\right) \beta_k k_{it} + \left(\frac{1}{\eta_i} - 1\right) \beta_l l_{it} \\ & + \left(\frac{1}{\eta_i} \beta_m - \beta_m + 1\right) m_{it} - h(A_{it}) - \ln(1 - \tau_{it}) \end{aligned} \right] \quad (6)$$

τ_{it} 为销售税率。由于地区间税收优惠政策和税收努力程度的差别,这里设定税率是企业异质的。本文的任务是根据式(2)、式(3)、式(5)和式(6),估计生产函数式(1)和需求函数式(4)的参数,进而识别出企业实际生产率 ω_{it} 和期望生产率(条件期望函数) $g(\omega_{it-1}, r_{it-1})$ 。

将式(2)和式(3)代入式(5),用 A_{it} 的三阶多项式逼近未知函数 $h(A_{it})$; 用 ω_{it-1} 的三阶多项式逼近未知函数 $g_{01}(\omega_{it-1})$; 用 ω_{it-1} 和 r_{it-1} 的三阶多项式逼近未知函数 $g_{11}(\omega_{it-1}, r_{it-1})$,复合误差项可以表示成:

$$\begin{aligned} \hat{\psi}_{it} &= rev_{it} - \left(1 - \frac{1}{\eta_i}\right) (\beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it}) - a_1 A_{it} - a_2 A_{it}^2 - a_3 A_{it}^3 - p_t \\ &\quad - c_0 - (c + \gamma_1 \cdot d_{07} + \gamma_2 \cdot d_{08}) \left(1 - \frac{1}{\eta_i}\right) - 1(R_{it-1} > 0) \\ &= 0 \left(1 - \frac{1}{\eta_i}\right) [g_{01}(\omega_{it-1})] - 1(R_{it-1} > 0) \left(1 - \frac{1}{\eta_i}\right) [g_{11}(\omega_{it-1}, r_{it-1})] \end{aligned} \quad (7)$$

我们用行业一般价格水平 p_t 代理总市场特征 φ_t 。 d_{07} 和 d_{08} 为时间虚拟变量。 c_0 吸收 $h(\cdot)$ 中的常数项; c 综合基年(2006)、 g_{00} 和未知函数 $g_{01}(\omega_{it-1})$ 中的常数项; $g_1 = g_{10} - g_{00}$ 。根据式(6):

$$\hat{\omega}_{it-1} = \frac{\eta_i}{\eta_i - 1} \left[\begin{aligned} & p_{M_{t-1}} - p_{t-1} - \ln \left(1 - \frac{1}{\eta_i}\right) + \left(\frac{1}{\eta_i} - 1\right) \beta_k k_{it-1} + \left(\frac{1}{\eta_i} - 1\right) \beta_l l_{it-1} \\ & + \left(\frac{1}{\eta_i} \beta_m - \beta_m + 1\right) m_{it-1} - a_1 A_{it-1} - a_2 A_{it-1}^2 - a_3 A_{it-1}^3 - \ln(1 - \tau_{it}) \end{aligned} \right] \quad (8)$$

$\ln \beta_m$ 和 $h(\cdot)$ 中的常数项被吸收到式(7)中的对应部分。式(7)和式(8)构成本文的基本估计方程。给定待估参数向量 θ ,可以得到式(7)的残差为 $\hat{\psi}_{it}(\theta)$ 。这样参数估计的 GMM 问题为:

$$\min_{\theta} \left[\frac{1}{N} \sum_{i,t} Z_{it} \hat{\psi}_{it}(\theta) \right]' W_N \left[\frac{1}{N} \sum_{i,t} Z_{it} \hat{\psi}_{it}(\theta) \right] \quad (9)$$

W_N 为权重矩阵, N 为总观测数。我们运用 Hansen(1982)的二阶段 GMM 估计。第一阶段使用权重矩阵 $W_N = \left[\frac{1}{N} \sum_{i,t} Z_{it} Z_{it}' \right]^{-1}$,得到 θ 的一致估计 θ ,以及 $\psi_{it}(\theta)$ 。第二阶段使用权重矩阵 $W_N = \left[\frac{1}{N} \sum_{i,t} Z_{it} \psi_{it}(\theta) \psi_{it}(\theta)' Z_{it}' \right]^{-1}$,得到最优估计 $\hat{\theta}$ 。

式(7)和式(8)中需要估计的参数有 23 个:3 个常数(c_0, c 和 g_{00});2 个生产率演化的时间趋势(γ_1 和 γ_2);3 个生产函数参数(β_k, β_l 和 β_m);未知函数 $h(\cdot)$ 的 3 个参数(a_1, a_2 和 a_3);未知函数 $g_{01}(\cdot)$ 的 3 个参数;以及未知函数 $g_{11}(\cdot)$ 的 9 个参数。运用 GMM 估计时, θ 中并不需要包括所有 23 个参数,根据 Doraszelski 和 Jaumandreu(2013)提出的“Concentrate out”方法,从式(7)和式(8)可见 θ 中只需要包括 $\beta_k, \beta_l, \beta_m, a_1, a_2$ 和 a_3 ,其他 17 个参数可以用这 6 个参数线性表示。这大大减少了 GMM 估计过程中非线性的程度,提高了算法的搜索效率。

根据 Newey(1990,1993)的最优工具理论,我们使用的工具变量共 47 个,包括常数,时间虚拟变量(2 个),销售努力的三阶多项式(滞后 1 期,3 个),R&D 投入量(对数)的三阶多项式(滞后 1 期,3 个),有 R&D 投入(虚拟变量)与资本、劳动和中间材料完全三阶多项式的交互(滞后 1 期,19 个),无 R&D 投入(虚拟变量)与资本、劳动和中间材料完全三阶多项式的交互(滞后 1 期,19 个)。

三、估计结果

本文所用数据来自国家统计局“全部国有和规模以上非国有工业企业数据库”,时间涉及 1998—2008 年。参照 Brandt 等(2012)的研究,本文对这套数据库进行了细致的匹配,尽量构建合理的面板(整理过程参见尹恒和杨龙见(2014))。^① 由于只有 2001 年、2005—2007 年存在 R&D 投入数据,而估计过程中需要用到滞后一期的变量,所以本文最终使用的样本期为 2005—2008 年,包括所有在此期间至少存在连续两年完整数据的企业。本文的行业定义参考 Doraszelski 和 Jaumandreu(2013),将二位数行业合并为 10 个行业。^②

1. 参数估计结果和模型设定检验

以上理论模型和结构设定是否符合中国制度背景?或者说这些估计设定能否较好地契合中国特定时期的数据产生过程?这是事关 R&D 投入回报估计结果可靠性和可信性的关键问题。为此我们先讨论基本参数估计结果的合理性,并根据 GMM 估计的特性检验模型的关键设定。

表 1 报告了所有 10 个行业的生产函数和需求函数的参数估计。可以看到,各行业资本、劳动和中间材料的弹性均处在合理的范围内,且都相当显著。长期规模收益(资本、劳动和中间材料弹性之和)在 0.9—1.1 之间,短期规模收益(劳动和中间材料弹性之和)在 0.9 左右,与文献的标准结果是一致的(Doraszelski 和 Jaumandreu,2013)。各行业的平均溢价在 1.1 左右,这是文献中溢价的标准估计结果(Aw 等,2011;De Loecker 和 Warzynski,2012)。大部分行业销售努力的一次项显著为正,考虑到销售努力处于 0—1 之间,可以看到,销售努力对需求的边际效应也是正的,这表明销售努力的加大确实能够使企业的需求曲线向外移。

表 1 参数估计结果

行业	生产函数			销售努力			溢价	函数值
	资本	劳动	材料	1 次项	2 次项	3 次项		
1	0.146 (0.009)	0.232 (0.023)	0.619 (0.022)	4.908 (2.845)	1.054 (29.864)	0.227 (85.784)	1.121	322.717
2	0.129 (0.005)	0.140 (0.022)	0.729 (0.022)	9.760 (3.372)	-13.982 (74.239)	-4.126 (291.279)	1.080	197.809
3	0.130 (0.013)	0.114 (0.037)	0.768 (0.035)	6.015 (4.741)	-14.767 (70.621)	-3.333 (210.222)	1.104	42.983

①由于篇幅限制,所用变量的定义与构建未列出,如有需要可向作者索要。

②由于篇幅限制,行业对应代码、主要变量的统计描述以及相应的附录均未列出,如有需要可以向作者索要。

续表 1 参数估计结果

行业	生产函数			销售努力			溢价	函数值
	资本	劳动	材料	1 次项	2 次项	3 次项		
4	0.107 (0.015)	0.207 (0.040)	0.705 (0.029)	3.839 (4.147)	15.100 (81.199)	3.677 (270.965)	1.107	111.607
5	0.160 (0.009)	0.343 (0.022)	0.580 (0.026)	-57.770 (3.007)	469.836 (35.826)	-857.805 (104.655)	1.126	185.631
6	0.106 (0.008)	0.100 (0.018)	0.739 (0.020)	11.521 (3.409)	-61.319 (48.920)	101.264 (128.926)	1.110	165.501
7	0.061 (0.008)	0.560 (0.020)	0.477 (0.017)	-10.851 (5.173)	22.377 (81.538)	7.738 (317.885)	1.114	513.875
8	0.107 (0.007)	0.096 (0.015)	0.819 (0.015)	4.497 (1.220)	-3.900 (22.657)	-0.881 (83.124)	1.117	123.523
9	0.134 (0.013)	0.103 (0.031)	0.786 (0.029)	6.922 (2.836)	-6.428 (50.459)	-2.544 (167.712)	1.105	61.387
10	0.141 (0.009)	0.118 (0.024)	0.795 (0.022)	-1.332 (2.328)	61.837 (42.011)	6.951 (160.179)	1.118	152.102

注:1.括号中为标准差;2.溢价为行业内企业的简单平均(未加权)。

下面检验是否有必要在模型中考虑 R&D 投入对期望生产率演化的影响,即模型的设定式(2)和式(3)是否能得到数据的支持。这是本文 R&D 投入回报估计模型的关键设定。如果企业生产率演化的真实过程如下:

$$\omega_{it} = g(\omega_{it-1}) + \xi_{it} \quad (10)$$

其中,R&D 投入并不影响企业生产率演化的马尔科夫过程,那么整个论文的分析思路就不成立。反之,如果真实情况是式(2)和式(3),但在模型误设为式(10)时,R&D 投入对下一期生产率的影响将进入 ξ_{it} 项。此时在结构估计中即便使用滞后期变量为工具,残差项中包含的滞后期的 R&D 投入也很可能与滞后期的工具变量相关,导致矩条件 $E[z_{it}\xi_{it}] = 0$ 失效,从而导致估计偏误。为此我们对基本模型进行参数约束检验。无约束模型为考虑 R&D 投入的式(2),有约束模型为式(10),即 R&D 投入对应的参数都为 0。根据 GMM 的基本理论,有约束模型和无约束模型采用同一个最优权重矩阵得到的最小化目标函数值之差,经过样本量的调整后服从开方分布,其自由度就是约束的个数。原假设是式(10)成立。据表 2 所示,所有行业最小化的目标函数值之差是显著的,这表明原假设被拒绝,R&D 投入对企业生产率的演化存在显著影响,即式(2)和式(3)设定是合理的。

表 2 R&D 投入对企业生产率演化影响的检验

行业	1	2	3	4	5	6	7	8	9	10
χ^2	172.864	62.500	18.960	28.646	331.827	52.745	36.300	43.209	20.276	265.044
p 值	0.000	0.000	0.041	0.001	0.000	0.000	0.000	0.000	0.027	0.000

注:开方分布的自由度为 10。

估计基本模型时,我们用三阶多项式来逼近期望生产率函数,以考虑 R&D 投入对期望生产率影响的非线性。即设定:

$$g_{11}(\omega_{it-1}, r_{it-1}) = \alpha_1 \omega_{it-1} + \alpha_2 \omega_{it-1}^2 + \alpha_3 \omega_{it-1}^3 + \alpha_4 r_{it-1} + \alpha_5 r_{it-1}^2 + \alpha_6 r_{it-1}^3 + \alpha_7 \omega_{it-1} r_{it-1} + \alpha_8 \omega_{it-1}^2 r_{it-1} + \alpha_9 \omega_{it-1} r_{it-1}^2 \quad (11)$$

这种非线性的逼近是否必要?或者说 R&D 投入对期望生产率影响是否具有非线性?具体地,若原假设:

$$\alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = \alpha_9 = 0 \quad (12)$$

成立,则 R&D 投入对期望生产率的影响是线性的。这也是一个线性约束检验,我们用同样的方法检验,结果见表 3。可以看到,10 个制造业行业中有 8 个十分显著,原假设被拒绝,

R&D 投入对期望生产率的影响是非线性的。值得注意的是行业 3(木材和家具制造业)不能拒绝 R&D 投入对期望生产率的影响主要是线性的。^①

表 3 R&D 投入对企业生产率影响非线性的检验

行业	1	2	3	4	5	6	7	8	9	10
χ^2	150.598	36.502	1.536	13.021	46.037	27.477	28.507	18.822	9.368	224.865
p 值	0.000	0.000	0.957	0.043	0.000	0.000	0.000	0.005	0.154	0.000

注:开方分布的自由度为 6。

在式(2)中 ξ_{it} 代表企业生产率演化过程中面临的随机冲击,我们可以进一步用 ξ_{it} 方差与实际生产率方差之比 $\frac{\text{var}(\xi_{it})}{\text{var}(\omega_{it})}$ 衡量 R&D 投入影响的不确定性。表 4 的结果表明,对于所有行业,在生产率演化路径中企业不可控制的随机扰动 ξ_{it} 占整个生产率变动的比重不可忽视,R&D 投入对企业实际生产率的影响确实存在很强的不确定性。即便对于不能拒绝线性影响的行业 3(木材和家具制造业),这种不确定性也是明显的。

表 4 企业生产率演化的不确定性

行业	1	2	3	4	5	6	7	8	9	10
$\frac{\text{var}(\xi_{it})}{\text{var}(\omega_{it})}$	0.503	0.345	0.192	0.181	0.471	0.273	0.336	0.350	0.216	0.865

总之,参数估计结果都在合理的范围内,检验结果支持我们关于 R&D 投入影响生产率演化机制的设定,也支持这种影响的不确定性和非线性的设定。这为我们以下的分析提供了可靠的基础。

2. R&D 投入对企业生产率分布的影响

下面讨论总体 R&D 投入对企业生产率的影响。表 5 列示了研发企业和无研发企业生产率的均值和方差差异。表 5 左半部分显示,平均而言,所有行业上期进行 R&D 的企业都具有更高的生产率。 t 检验表明两组企业的差距十分显著。表 5 右半部分表明,所有行业进行 R&D 的企业下期生产率的标准差更大, F 检验显示差距十分显著。这进一步说明了 R&D 投入回报的不确定性。

表 5 R&D 投入与下期生产率均值和标准差

行业	生产率均值差异检验				生产率标准差差异检验			
	无研发 均值	有研发 均值	有研发企业 TFP 均值更高		无研发 标准差	有研发 标准差	有研发企业 TFP 标准差更高	
			t 值	p 值			F 值	p 值
1	-0.022	0.202	13.873	0.000	1.022	1.113	1.187	0.000
2	-0.013	0.206	18.685	0.000	0.778	0.842	1.172	0.000
3	-0.010	0.196	6.957	0.000	0.904	0.964	1.138	0.010
4	-0.028	0.511	15.170	0.000	1.033	1.314	1.620	0.000
5	-0.092	0.592	49.966	0.000	0.982	1.446	1.396	0.000
6	-0.019	0.225	12.187	0.000	0.955	1.168	1.496	0.000
7	-0.022	0.296	16.205	0.000	1.024	1.201	1.375	0.000
8	-0.040	0.240	34.844	0.000	0.610	0.799	1.717	0.000
9	-0.064	0.305	21.316	0.000	0.903	1.119	1.536	0.000
10	-0.116	0.423	39.845	0.000	1.065	1.475	1.918	0.000

进一步运用 Kolmogorov-Smirnov(KS)检验讨论总体上 R&D 投入对企业生产率的演化是否具有促进作用。作为一种非参数检验方法,KS 检验可以确定两组样本量的分布是否显

^①我们仍然采用设定式(11)对这个行业进行估计。即使原假设成立,用设定式(11)进行估计只是扩大了估计参数集,参数估计的一致性不受影响。

著相同。该检验不依赖均值的位置,具有较强的稳健性。具体地,设 $G_{N_1}(\cdot)$ 和 $F_{N_0}(\cdot)$ 分别是有研发和无研发企业下期生产率的分布函数, N_1 和 N_0 分别是有无研发的企业数目。双边检验的原假设是两组企业生产率的分布相同,也即:

$$H_0: G_{N_1}(\bar{g}) - F_{N_0}(\bar{g}) = 0 \quad \forall \bar{g} \quad (13)$$

单边检验的原假设是有研发投入的企业比没有研发投入的企业在分布上随机占优,也即:

$$H_0: G_{N_1}(\bar{g}) \leq F_{N_0}(\bar{g}) \quad \forall \bar{g} \quad (14)$$

检验统计量分别为:

$$S^1 = \sqrt{\frac{N_0 N_1}{N_0 + N_1}} \max_{\bar{g}} \{ |G_{N_1}(\bar{g}) - F_{N_0}(\bar{g})| \} \quad (15)$$

$$S^2 = \sqrt{\frac{N_0 N_1}{N_0 + N_1}} \max_{\bar{g}} \{ G_{N_1}(\bar{g}) - F_{N_0}(\bar{g}) \} \quad (16)$$

概率值的计算通过极限分布得到:

$$P(S^1 > c) = -2 \sum_{k=1}^{\infty} (-1)^k \exp(-2k^2 c^2) \\ P(S^2 > c) = \exp(-2c^2) \quad (17)$$

据表 6 左半部分显示,所有行业中有研发和无研发这两类企业的分布显著不同,R&D 投入会改变下期生产率的分布。表 6 右半部分的检验结果显示,所有行业我们都不能拒绝有研发企业在下期生产率分布上一阶占优的假设,即在生产率的分布上,有 R&D 投入的企业存在显著的优势。R&D 投入对下期企业生产率存在显著的促进作用。

表 6 研发投入对企业 TFP 分布的影响

行业	Kolmogorov-Smirnov 检验			
	两分布比较		有研发的企业是否一阶占优	
	s1 值	p 值	s2 值	p 值
1	7.317	0.000	0.021	0.999
2	10.183	0.000	0.015	1.000
3	4.111	0.000	0.045	0.996
4	7.476	0.000	0.014	1.000
5	22.310	0.000	0.033	0.998
6	4.505	0.000	0.044	0.996
7	7.152	0.000	0.011	1.000
8	18.354	0.000	0.534	0.566
9	11.470	0.000	0.039	0.997
10	21.520	0.000	1.037	0.116

3. R&D 投入弹性的异质性

前面的讨论表明,本文估计企业 R&D 投入回报的理论模型和结构设定是合理的,从总体上看 R&D 投入对企业生产率的演化具有积极影响。下面我们进一步分析这种影响的企业异质性。估计出了预期生产率函数 $g(\omega_{it-1}, r_{it-1})$ 后,就可以直接估计 R&D 投入的边际弹性(只针对有研发投入的企业)。如前文指出, $\frac{\partial g(\omega_{it-1}, r_{it-1})}{\partial r_{it-1}}$ 即为 $t-1$ 期 R&D 投入对 t 期预期产出的边际弹性,可以用来度量 R&D 投入的有效性。表 7 和图 1 列示了 R&D 投入边际弹性的分布情况。首先,从均值上看,R&D 投入边际弹性与现有文献对中国制造业的估计相差并不大,平均意义上 R&D 投入的回报很低,甚至为负(如行业 10)。其次,R&D 投入回报确实具有很强的企业异质性。行业 1(食品、饮料和烟草)和行业 10(电子) R&D 投入边际弹性的方差最大,接近于 1。如行业 10,R&D 投入边际弹性的第 25 分位值

为-0.332,而第75分位值达到0.355。 $R\&D$ 投入边际弹性为负也是合理的,边际上负的回报与全局上 $R\&D$ 投入的正向影响并不矛盾。企业可能会出于先发优势或者 $R\&D$ 投入完整性的考虑,导致短期内 $R\&D$ 投入的边际弹性为负。估计分布与前面的结果也是一致的。例如对于不能拒绝 $R\&D$ 投入线性影响的行业3,图1中其分布相对而言十分集中,表7中其方差也最小(只有0.06),即 $R\&D$ 投入边际弹性的企业异质性最小。

表7 预期生产率— $R\&D$ 投入边际弹性的分布特征

行业	q1	q2	q3	q4	q5	简单均值	加权均值	方差	偏度
1	-0.714	-0.371	-0.140	0.279	1.043	0.037	0.035	0.888	0.199
2	-0.093	-0.002	0.039	0.074	0.140	0.012	0.039	0.218	-0.125
3	-0.027	-0.006	0.003	0.026	0.063	0.007	0.013	0.060	0.060
4	-0.125	-0.074	-0.016	0.047	0.145	0.004	0.004	0.124	0.158
5	-0.257	-0.114	-0.034	0.121	0.370	0.016	0.002	0.320	0.158
6	-0.192	0.010	0.067	0.143	0.227	0.033	0.045	0.235	-0.145
7	-0.131	-0.032	0.011	0.058	0.141	0.005	-0.011	0.147	-0.043
8	-0.126	-0.010	0.037	0.101	0.182	0.029	0.030	0.183	-0.042
9	-0.047	-0.025	-0.007	0.036	0.098	0.012	0.026	0.078	0.241
10	-0.845	-0.332	0.065	0.355	0.854	-0.041	0.034	1.066	-0.09

注:1.q1、q2、q3、q4和q5分别表示第10分位、第25分位、50分位、75分位和90分位值。2.计算 $R\&D$ 投入边际弹性的加权均值时删除了生产率上下5%,以销售收入为权重。3.偏度定义为:(均值-中位数)/标准差。

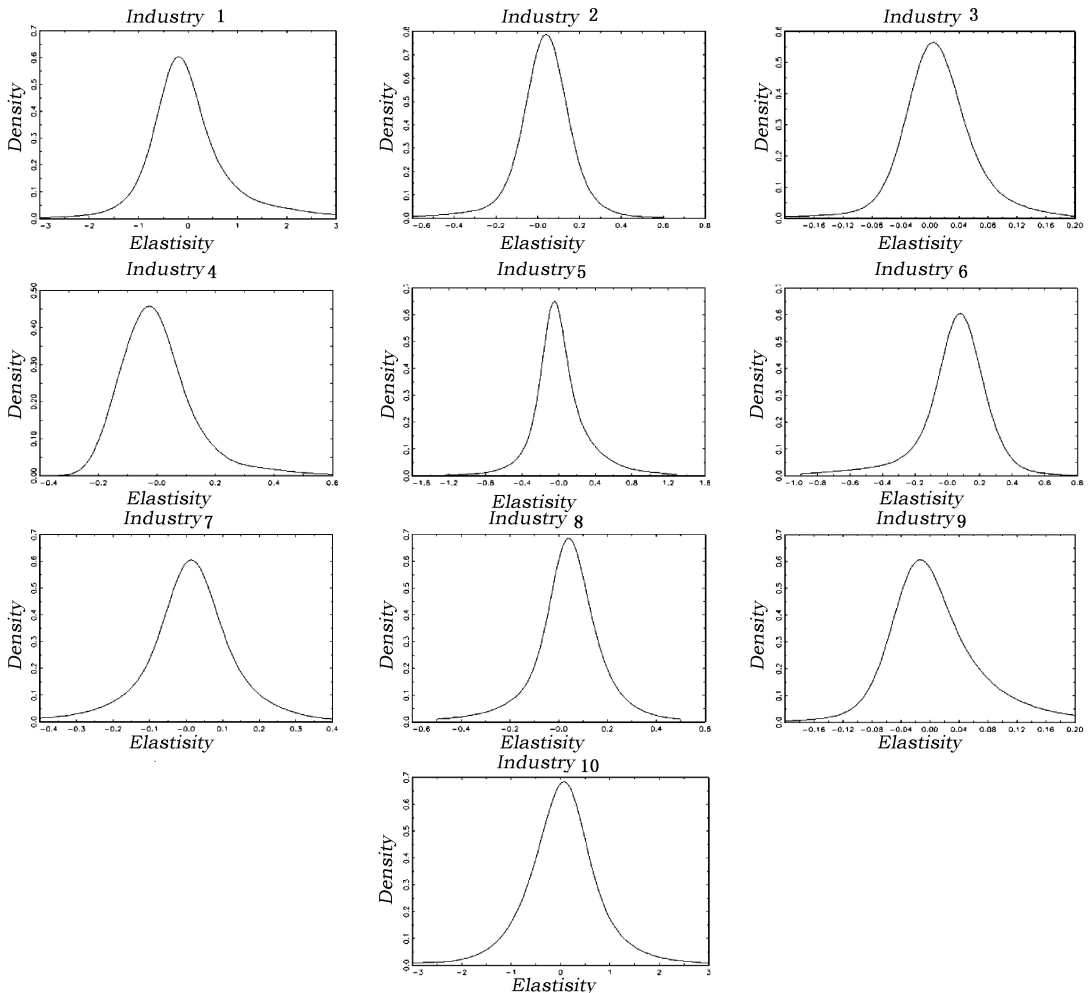


图1 预期生产率— $R\&D$ 投入边际弹性的分布图

我们还借鉴文献中的传统做法,对有研发投入的企业样本用滞后一期的 R&D 投入回归估计的企业生产率,得到平均的 R&D 投入弹性(本文称为 R&D 投入平均弹性),如表 8 所示。可以看出,平均意义上所有 10 个行业的 R&D 投入都对下期实际产出存在显著促进作用。而且,这里估计的 R&D 投入平均弹性值与国内外现有文献的估计基本一致。这在一定程度上也表明 R&D 投入的平均回报确实处于一个比较小的区域。

表 8 实际生产率—R&D 投入平均弹性

行业	1	2	3	4	5	6	7	8	9	10
系数	0.023	0.020	0.020	0.049	0.056	0.024	0.029	0.025	0.029	0.044
t 值	18.37	22.13	8.52	20.60	57.21	17.11	20.99	48.72	24.44	51.00

注:其他控制变量包括:时间虚拟,地区虚拟以及企业控股状况虚拟。

四、总结性评论

本文的结构估计模型将 R&D 投入引入企业生产率演化的马尔科夫过程,正式处理 R&D 投入对企业生产率影响的不确定性、非线性和企业异质性,并同时考虑企业需求的异质性。所有 10 个行业生产函数和需求函数的参数估计均处在合理的范围内,参数约束检验也表明基本模型的设定是合理的,R&D 投入对企业生产率的演化存在显著影响。这说明使用该模型评估 R&D 投入的产出绩效是可靠的。估计结果表明,R&D 投入对其下期生产率的影响确实具有不可忽视的不确定性、非线性和企业异质性。所有行业 R&D 投入会显著改变行业内企业的生产率分布,进行 R&D 投入的企业下期生产率分布显著优于没有进行 R&D 投入的企业。在充分考虑了影响的不确定性、非线性和企业异质性之后,中国制造业所有行业的 R&D 投入都能够显著地促进企业生产率发展。

在平均意义上,本文估计的各行业 R&D 投入平均弹性及 R&D 投入边际弹性,与使用知识资本模型的国内外文献的估计值并不存在明显差距。然而,所有行业中企业 R&D 投入的回报确实都具有很强的企业异质性。也正是这种异质性为理解 R&D 投入回报之谜,即平均意义上 R&D 投入的回报较低与宏观层面 R&D 投入对经济增长的重要性共存问题提供了钥匙。即使平均而言 R&D 投入的回报并不高,但由于其具有高度的企业异质性,一些企业仍然存在较强的激励进行 R&D 投入;又由于 R&D 投入取得的新工艺和新产品具有很强的外溢性,部分企业由高额预期回报驱动的 R&D 投入活动可以成为推动整个经济增长的重要动力。中国经济和社会发展正进入一个全新的时期,人口红利在逐渐消失,资本数量上的扩张也不能再成为经济增长的主要推动力。我们有充分的理由相信,在新阶段 R&D 投入及由此带来的生产率的改善能够成为推动中国经济持续发展的主要力量。

当然,R&D 投入是影响企业长期发展的动态决策。受到数据限制,本文只对四年 R&D 投入回报进行了估计。随着数据可获得性的增强,本文提出的估计思路可以纳入更为完整的企业动态决策结构,以利用更为丰富的面板信息。

主要参考文献:

- [1]李小平,朱钟棣.国际贸易、R&D 溢出和生产率增长[J].经济研究,2006,(2):31—43.
- [2]陆国庆.中国中小板上市公司产业创新的绩效研究[J].经济研究,2011,(2):138—148.
- [3]毛德凤,李静,彭飞,等.研发投入与企业全要素生产率——基于 PSM 和 GPS 的检验[J].财经研究,2013,(4):134—144.
- [4]吴延兵.R&D 与生产率——基于中国制造业的实证研究[J].经济研究,2006,(11):60—71.

- [5]尹恒,杨龙见.地方财政对本地居民偏好的回应性研究[J].中国社会科学,2014,(5):96—115.
- [6]张海洋.R&D两面性、外资活动与中国工业生产率增长[J].经济研究,2005,(5):107—117.
- [7]周亚虹,贺小丹,沈瑶.中国工业企业自主创新的影响因素和产出绩效研究[J].经济研究,2012,(5):107—119.
- [8]宗庆庆,周亚虹.内生情形下企业研发对生产率作用评估[J].世界经济文汇,2013,(6):39—54.
- [9]Doraszelski U, Jaumandreu J. R&D and productivity: Estimating endogenous productivity[J]. Review of Economic Studies, 2013, 80(4): 1338—1383.

Understanding the Mystery of R&D Investment Returns: The Introduction of Uncertainty, Nonlinearity and Firm Heterogeneity

Yin Heng¹, Liu Di²

(1. National Academy of Development and Strategy, Renmin University of China, Beijing 100872, China; 2. Business School, Beijing Normal University, Beijing 100875, China)

Abstract: The existing research finds that R&D investment at firm level has low or even negative performance. And it cannot support the inference that R&D investment has important effect on economic growth. The reason lies in that traditional research based on knowledge capital model estimates average returns of R&D investment and smooths out the firm heterogeneity of returns of R&D investment. This paper constructs a structural model to deal with uncertainty, nonlinearity and heterogeneity of returns of R&D investment. Compared with existing structural models, it takes firm heterogeneity of premium into account, and specifically addresses factors affecting firm demand. It arrives at the conclusions as follows: firstly, the returns of R&D investment are actually uncertain and nonlinear, thereby having very obvious firm heterogeneity; secondly, the returns of R&D investment in individual enterprises can even be higher than average returns several times, which is enough to provide a strong incentives for R&D investment. Firm heterogeneity of returns of R&D investment is the key to the understanding of coexisting of low returns of R&D investment in average sense and the importance of R&D investment to economic growth at macro level. Therefore, it fully believes that R&D investment and resulting productivity improvement can be the main force for promoting sustainable economic development in China in the new normal.

Key words: return of R&D investment; firm productivity; firm heterogeneity; structural estimation

(责任编辑 石头)