

工业增加值与全要素生产率估计

——基于中国制造业的拟蒙特卡洛实验

朱沛华, 陈林

[摘要] 全要素生产率是现代化经济体系的主要统计指标之一,但其存在诸多统计偏差问题有待克服。其中,由工业增加值代理产出项(相对总产值而言)而导致的估计问题尤为严重。为此,本文从数理模型上分析增加值偏差对于全要素生产率估计结果的影响;在实证检验上分别构建总产值与增加值生产函数、总产值与增加值成本函数,并使用半参数估计法进行估计。拟蒙特卡洛实验的结果表明,由于增加值偏差的存在,增加值生产函数测算得到的全要素生产率异质性要大于总产值生产函数,增加值成本函数测算得到的全要素生产率异质性要大于总产值成本函数,使用总产值作为产出代理变量得到的估计结果更为稳健。以1999—2013年中国工业企业数据为样本的检验结果表明,2011—2013年比2008年以前全要素生产率分布的所有制差异明显缩小;其动态分解结果表明,2011—2013年全要素生产率变动出现停滞甚至下滑的趋势,不同行业的全要素生产率变化表现出明显的差异。在后续研究中需要关注全要素生产率在分布上的新特征以及变动上的新趋势,以期得到更为贴近现实的结论。

[关键词] 全要素生产率; 工业增加值; 全要素生产率估算偏差; 增加值偏差; 成本函数

[中图分类号]F424 [文献标识码]A [文章编号]1006-480X(2020)07-0024-19

一、引言

变革,一直是改革开放以来中国经济发展的“主题词”。2017年10月,党的十九大报告展示了深化改革的决心——“我国经济已由高速增长阶段转向高质量发展阶段,正处在转变发展方式、优化经济结构、转换增长动力的攻关期”“推动经济发展质量变革、效率变革、动力变革,提高全要素生产率”。不难看出,在党的十九大报告第五部分“贯彻新发展理念,建设现代化经济体系”中,质量与效率是统领全章的关键词。那么,如何考察这次“攻关期”变革的绩效?如何衡量经济增长的质量、经

[收稿日期] 2019-12-01

[基金项目] 教育部哲学社会科学研究重大课题攻关项目“我国全要素生产率提升与测算研究”(批准号17JZD013);研究阐释党的十九届四中全会精神国家社会科学基金重点项目“自由贸易试验区负面清单制度的推广与评估”(批准号20AZD050);国家自然科学基金面上项目“竞争政策与准入规制的协调机制研究”(批准号71773039)。

[作者简介] 朱沛华,中国人民大学经济学院博士研究生;陈林,暨南大学产业经济研究院教授,博士生导师,经济学博士。通讯作者:陈林,电子邮箱:charlielinchen@qq.com。感谢广东省高校“珠江学者”岗位计划的资助和中国人民大学李靖博士、孟祥慧博士的有益评价,感谢匿名评审专家和编辑部的宝贵意见,当然文责自负。

济发展的效率?本文认为,答案正在党的十九大报告第五部分中,即“提高全要素生产率”。作为当前国内经济学实证研究的主流领域之一,“全要素生产率”(Total Factor Productivity,简称TFP)首次进入了最高级别的党和国家公文,这从侧面体现出中央对今后变革经济发展方式的决心。科学、客观地度量全要素生产率,是实务界、理论界亟需解决的重大问题,是新时代衡量全面深化改革绩效的重要工作。

TFP一直是衡量工业生产投入产出效率的核心指标,但其一直缺乏完全达成学术共识的衡量标准(Syverson,2011)。从函数的设定形式看,常见的函数形式可以是科布—道格拉斯(C-D)函数或超越对数(Translog)函数。在具体的估计方法上,也存在参数法(OLS、固定效应等)、非参数法(数据包络方法等)以及半参数估计法——OP方法(Olley and Pakes,1996)、LP方法(Levinsohn and Petrin,2003)、ACF方法(Ackerberg et al.,2015)等。然而,由于各类TFP估算方法均存在局限性(于永达和吕冰洋,2010;柳荻和尹恒,2015),以往实证研究检验出来的TFP千差万别,甚至存在不少相互矛盾的估计结果。若不能够准确地测度TFP,有效识别各类估算方法潜在的统计问题,那么得到的结论必然背离现实,这也是导致现有研究存在分歧的重要原因。

“增加值偏差”(Value-Added Bias)问题扮演了重要角色——直接影响着TFP的估计信度。在生产函数与成本函数的常用技术设定中,主要存在总产值(Gross-Output)与增加值(Value-Added)两种核算方法。直觉上,这两种技术设定似乎是可以相互替代的。然而,在实际估计过程中,两类方法得到的TFP不尽相同,增加值函数的估计存在着明显的偏差,最终使得测算得到的TFP离散程度更大。所谓增加值偏差,是指增加值类型的生产(成本)函数由于没有考虑到中间投入(价格)因素所引致的一系列偏差(Gandhi et al.,2016)。具体体现在:①增加值作为产出变量并没有考虑到企业中间投入份额变化对生产函数估计的影响;②中间投入能够显著地影响企业的产出效率,缺乏中间投入会造成一定程度的遗漏变量问题。除此之外,遗漏其他不可测的生产要素变量引起的估计偏差是另一个常见的计量问题,计量模型设置的不准确对生产函数与成本函数的估计也会造成显著的差异。存在遗漏生产要素的情况下,这些变量的影响将会被错误地识别为TFP波动,使得测算结果产生明显的偏差,并与实际TFP产生偏离,研究结论的可信度将大打折扣。本文着重研究由于增加值偏差与遗漏不可观测变量引起的TFP估计偏误,比较同等条件下生产函数与成本函数估计结果的差异。

在已有研究的基础上,本文可能的边际贡献在于:①立足于中国特殊情境,在Gandhi et al.(2016)提出的增加值偏差基础上展开,以中国工业企业数据库为样本,构建拟蒙特卡洛实验,检验以增加值和总产值分别作为产出项对生产函数与成本函数估计的影响以及估算得到TFP分布的差异,并进一步比较了未辨明的投入要素对TFP分布的影响,为研究中国制造业的生产效率提供有益的理论建议。②对中国工业企业数据2011—2013年样本中间投入、增加值等关键缺失指标进行估算,研究制造业企业的TFP分布特征以及动态变化,分析由于增加值偏差造成的结论性偏差,并探讨美国次贷危机冲击后中国制造业TFP出现的新特征与新变化。研究发现,2011—2013年相比于2008年以前中国工业企业的TFP分布呈现出所有制差异明显缩小,各类子行业层面的TFP改进呈现出分化等新特征,为制造业TFP的估计与探讨提供新的经验证据。

剩余部分的结构安排如下:第二部分为文献回顾;第三部分为理论分析;第四部分为实证方法;第五部分为模拟分析;第六部分为实证分析;最后为主要结论与启示。

二、文献回顾

近年来,全要素生产率的相关领域研究越来越受到学界的重视。以中国知网收录的文献为例,1999—2008年,国内学界平均每年发表125篇全要素生产率相关的期刊论文;2009—2018年,该指标上升至每年837篇。^①2018年发表了1193篇相关的论文,是1999年发文量的27.1倍(44篇)。从发文质量看,2018年相对高质量的CSSCI和北大核心期刊的论文占比59.5%(710篇),发文量是1999年的33.8倍(21篇)。

尽管关于TFP的研究如此丰富,但不同文献对同一样本测算得到的TFP千差万别,在实证检验上不乏矛盾之处(鲁晓东和连玉君,2012;杨汝岱,2015)。其中,关于国有企业TFP的分歧最为突出。比如,国有部门的资源配置是否存在效率改进?对该问题,黄先海等(2017)使用增加值生产函数估算1998—2007年制造业企业的全要素生产率,研究发现“抓大放小”的改革政策既能提高国有部门内部的要素配置效应,也能促进国有部门间的要素流动,最终提高了制造业的TFP。然而,有文献给出了相反的答案。聂辉华和贾瑞雪(2011)使用总产值生产函数估算了TFP,发现1999—2007年国有企业的资源配置效率低下,进入和退出行为没有发挥明显的作用。王永进和刘灿雷(2016)发现在价值链上游,由于国有企业的垄断,低效率国有企业滞留市场中,技术进步与资源配置效应受到抑制。

又比如,国有企业的效率是否存在着赶超趋势?持肯定观点的文献认为国有企业表现出效率赶超的趋势。周黎安等(2007)基于增加值产出的LP估计方法测算了1998—2004年制造业企业的TFP,发现国有企业相对其他所有制企业TFP较低,却呈现明显的追赶效应。然而,持否定观点的文献认为国有企业的效率改进速度处于劣势。Brandt et al.(2008)使用同期的研究数据发现,国有企业的TFP虽然呈现上升趋势,却仍远低于同期的私有企业效率。贺胜兵和周华蓉(2014)比较了2002—2011年工业企业中国有企业和民营企业的效率变动趋势,发现整体上国有企业的TFP增长率要低于民营企业。研究结论的不一致,固然源于研究样本选择差异、企业产权属性界定差异等现实因素,但不能忽略统计方法选择不当所造成的测量偏差。同步性、遗漏变量等均是使用索洛余值法测算TFP的过程中不可避免的统计问题,甚至选择刻画企业产出变量(总产值与增加值)的差异也会导致显著的TFP测算偏差。

TFP对于企业成长乃至经济社会持续发展的重要意义不言而喻。在实际运用中,由于技术设定的偏差,不能观测到或者进行量化的生产要素的影响,以及各类内生性问题的存在,使得实际测算的结果不尽如人意。已有文献针对TFP的来源设定、内涵外延以及各类测度方法进行比较,以找到最为合适的TFP测度方法(尹恒等,2015;张志强,2015;Akerberg et al.,2015)。Brandt et al.(2012)测度了中国制造业在进入WTO以来的TFP增长速度,发现1998—2007年增加值得到TFP平均增速为7.96%,总产值得到的TFP平均增速则是2.85%。Van Beveren(2012)在半参数估计法的基础上,构建了一个内生投入产出价值与产品选择的TFP估计模型,进而讨论了残差项离散程度、技术异质以及规模报酬等不同条件下的最优估计模型选择。Doraszelski and Jaumandreu(2013)用企业层面的产品价格与工资数据来拟合中间投入需求函数,并运用到TFP的估计框架当中,强调了企业知识资本积累对估算不可观测的TFP的重要意义。Grieco and McDevitt(2016)以医疗服务行业为例,指出医疗单位面临着增加诊断服务数量与提高服务质量的权衡,未考虑产出质量差异将导致TFP估计的失真。综上,现有的研究对于增加值与总产值的技术设定差异似乎并没有给予足够

^① 以期刊论文中主题或者篇名或者关键词包含“全要素生产率”进行检索得到的结果。

的重视,少有文献对二者的区别进行系统地论述分析。本文着重研究增加值偏差对于微观企业 TFP 异质性的影响,对相关方面的空白进行补充,为 TFP 的测算分析提供有益的建议。

三、理论分析

本节主要运用计量分析工具,对增加值偏差进行理论分析,参照 Gandhi et al.(2016)的分析方法,将增加值生产函数的估计偏差分解为同步性偏差与增加值偏差,证明增加值偏差的存在性。进一步地,对增加值成本函数的估计偏差进行分解,证明了该方法论下存在的同步性问题与增加值偏差。此外,针对遗漏不可观测变量问题以及可能造成的估计偏差进行简单分析,论述生产函数与成本函数在实证估计中的区别。依据理论分析结果,提出相应有待检验的理论假说,并运用拟蒙特卡罗实验进行验证。

1. 增加值生产函数的估计偏差

运用计量分析方法,证明增加值的生产函数与成本函数均存在着特殊的增加值偏差,最终会使得测算的全要素生产率异质性更大。设定总产值类型的生产函数, q 表示在生产函数技术设定下的理想产出, y 则表示在外生技术冲击下企业实际的产出, va 表示实际的增加值产出,为方便分析,假设企业产出的价格标准化为1。 ω 表示生产函数技术下滤除了外生技术冲击的 TFP, ε 表示外生技术冲击。 m 表示中间投入, P_m 表示中间投入价格, x 则是除了中间投入以外其余的可以观测的要素投入(例如,包含资本、劳动等投入要素)。

$$q=f(x,m) \cdot e^\omega \tag{1}$$

$$y=q \cdot e^\varepsilon \tag{2}$$

S 表示中间投入价值占总产出价值的份额。根据生产法,工业增加值为工业总产出与各项中间投入之差。由 $va=y-P_m m$ 可得式(3)。

$$va=y \cdot (1-S)=f(x,m) \cdot e^{\omega+\varepsilon} \cdot (1-S) \tag{3}$$

假设生产函数为 C-D 函数形式,对上式进行对数展开得到理想的计量方程式(4), α 与 α_m 表示对应要素投入与中间投入的待估系数。

$$\ln va \approx \ln x \alpha + \alpha_m \ln m + \omega + \varepsilon + \ln(1-S) \tag{4}$$

实际使用的增加值生产函数计量方程为式(5), e 为残差项。

$$\ln va = \ln x \alpha + e \tag{5}$$

实际得到的参数估计可以分解为式(6)。

$$\hat{\alpha} = \alpha + E[(\ln x' \ln x)^{-1} \ln x' \omega] + E[(\ln x' \ln x)^{-1} \ln x' (\alpha_m \ln m + \ln(1-S))] + E[(\ln x' \ln x)^{-1} \ln x' \varepsilon] \tag{6}$$

可知参数估计 $\hat{\alpha}$ 除了真实的 α 以外由三部分组成,由于外生冲击的假定可知右式第三部分为0,第一部分代表了同步性偏差(Simultaneous Bias),即 TFP 会影响当期的投入要素决策,这也是近年来半参数估计法(OP、LP、ACF等)致力于解决的内生性问题。右式第二部分即是增加值生产函数特有的偏差,主要是由于增加值生产函数将中间投入要素剔除到计量方程以外,因而使得估计方程不能够有效控制中间投入对于总产出的效应而引致的偏差。

即使不考虑同步性问题,最后测算的 TFP 也存在一定的偏差:

$$e - (\omega + \varepsilon) = [\ln x' (\alpha - \hat{\alpha}) + \alpha_m \ln m] + \ln(1-S) \tag{7}$$

式(7)等号右边第一部分代表了回归系数中未包含中间投入项而直接导致的偏差,第二部分是

由于生产函数未考虑到中间投入占总产出的份额而间接导致的偏差。由此可知,增加值生产函数相比总产值生产函数,除了未考虑到中间投入而导致的估计偏差,中间投入的份额变化也会导致 TFP 更大的异质性(Heterogeneity),测算得到的 TFP 离散度更高,由此提出:

假说 1:由于增加值偏差的存在,增加值生产函数的系数估计波动更大。

假说 2:由于增加值偏差的存在,增加值生产函数估计得到的 TFP 异质性更大。

2. 增加值成本函数的估计偏差

增加值成本函数的估计也存在着偏差,情况比生产函数更为复杂。设定总产值类型的成本函数如式(8)。 C 代表实际总成本, P_m 表示中间投入的要素价格, P_x 表示除中间投入外其他可以量化的要素投入价格(资本价格、劳动价格等)。 ω_c 表示成本函数技术下的剔除外生冲击后的 TFP, η 表示外生技术冲击。

$$C=f(y, P_x, P_m) \cdot e^{-\omega_c + \eta} \quad (8)$$

令 S_c 表示中间投入价值占总成本价值的份额,得到增加值的总成本表达式如式(9)所示:

$$C^{va} = C - P_m m = C \cdot (1 - S_c) = f(y, P_x, P_m) \cdot e^{-\omega_c + \eta} \cdot (1 - S_c) \quad (9)$$

假设成本函数为 C-D 形式,对上式进行对数展开得到理想的计量方程如式(10), β_y 、 β_m 与 β 表示待估系数。

$$\begin{aligned} \ln C^{va} &\approx \beta_y \ln y + \ln P_x \beta + \beta_m \ln P_m - \omega_c + \eta + \ln(1 - S_c) \\ &= \beta_y \ln va - \beta_y \ln(1 - S) + \ln P_x \beta + \beta_m \ln P_m - \omega_c + \eta + \ln(1 - S_c) \end{aligned} \quad (10)$$

实际使用的增加值生产函数计量方程如式(11), β_{va} 为待估参数, e_c 为残差项。

$$\ln C^{va} = \beta_{va} \ln va + \ln P_x \beta + e_c \quad (11)$$

实际的参数估计可以分解成式(12)。^①

$$\begin{aligned} (\widehat{\beta}_{va}, \widehat{\beta}')' &= (\beta_y, \beta')' - E[(X'X)^{-1} X' \omega_c] + \\ &E[(X'X)^{-1} X' (\beta_m \ln P_m - \beta_y \ln(1 - S) + \ln(1 - S_c))] + E[(X'X)^{-1} X' \eta] \end{aligned} \quad (12)$$

由上式可知,被解释变量的估计参数除了真实的 (β_y, β') 以外,由三部分组成。由于外生技术冲击的假设,第三部分 $E[(X'X)^{-1} X' \eta]$ 为 0。第一部分是同步性偏差问题,TFP 与要素价格弹性具有同期相关性。第二部分是增加值偏差,即增加值类型的成本函数由于没有考虑到中间投入价格因素的影响以及中间投入占总成本与总产出份额而导致的偏差。因此,即使不考虑同步性问题,使用增加值成本函数来测算 TFP 也存在着估计偏差,如式(13)所示。增加值偏差所带来的 TFP 估计偏误可以分为两部分:一是要素价格弹性中未考虑中间投入而直接导致的偏差;二是未考虑到中间投入占总产出与总成本份额的变化而间接导致的偏差。

$$e_c - (-\omega_c + \eta) = [X'((\beta_y, \beta')' - (\widehat{\beta}_{va}, \widehat{\beta}')') + \beta_m \ln P_m] + [\ln(1 - S_c) - \beta_y \ln(1 - S)] \quad (13)$$

与生产函数类似,由于增加值偏差的存在,使得该类型的成本函数估计系数波动性更大,最终得到的 TFP 自然也就具有更强的异质性,测算偏差也就随之被放大,由此提出:

假说 3:由于增加值偏差的存在,增加值成本函数的系数估计波动更大。

假说 4:由于增加值偏差的存在,增加值成本函数估计得到的 TFP 异质性更大。

^① 设定矩阵 $X=(\ln va, \ln P_x)$ 。

3. 遗漏不可观测变量引起的估计偏差

为了考察遗漏变量问题对函数估计的影响,引入不可观测的要素变量,将其置入到估计方程中进行比较。 z 表示不可观测或量化的投入要素, P_z 表示对应的要素价格,其余符号含义同上。^①

$$\ln y = \ln x \alpha + \varepsilon(z) \quad (14)$$

$$\ln C = \beta_y \ln y(x, z) + \ln P_x \beta + \eta(P_z) \quad (15)$$

可见,生产函数自身无法克服遗漏变量的问题,这源于某些特定的投入要素并不能通过量化的方法进入计量方程。从技术设定看,尽管成本函数的残差项也存在着遗漏要素价格的问题,但其自身包含实际产出项 $\ln y(x, z)$,能够识别不可量化生产要素的影响,最终一定程度上控制要素价格弹性的遗漏变量问题。以往的实证结果中,影响总成本的因素主要是产出项,各个要素价格弹性的估计系数较低,表明遗漏要素价格变量造成的估计偏差也较低,这使得成本函数的估计结果较为稳健,进一步提出:

假说 5:成本函数的遗漏变量问题更小,其估计结果相对生产函数更为稳健。

四、实证方法

1. 模型构建

为对以上的五个假说进行检验,本文选用常用的生产函数与成本函数作为计量模型进行研究。为了控制同步性问题对实证结果的影响,使用 Olley and Pakes(1996)提出的半参数估计方法对柯布—道格拉斯(C-D)生产函数进行估计。此外,同步性问题的存在使得成本函数的要素价格估计也是有偏的。参照该方法的核心思路,对成本函数进行半参数方法估计。基于生产函数的 OP 方法十分常用,在此不再赘述,主要就成本函数的半参数估计方法进行说明。

近年来,企业面临的产品需求异质性已经成为不可忽视的问题。^②由于统计核算问题,中国工业企业也存在着 TFP 的测算“困境”,即现有的统计口径通常披露企业的产值,而未核算企业的产出数量以及产品价格。据不完全统计,当前国内主要有两种做法来解决“产值问题”:①王永进和冯笑(2019)以及许明和李逸飞(2020)通过合并中国工业企业数据库与中国海关数据,以海关出口的单价指标构建价格指标,从而克服“产值问题”带来的影响。②陈林(2018)则使用了独特的行业样本——市政公用行业。由于这些行业的价格调整需要“价格听证会”,且所有大城市每年定价会在部分统计年鉴中公布,可以使用统计年鉴公布的年度数据对“产值问题”进行控制。然而,上述两类方法均会导致有效样本的剧减并降低实证结果的效度。

为此,本文主要参照 De Loecker(2011)、Bloom et al.(2013)以及尹恒和张子尧(2019)等研究的建模思路,引入行业层面的总需求因子 q_{st} 以及企业层面的横向产品差异因子 H_{it} 。其中,下标 s 为四位数行业代码,下标 i 代表企业个体代码,下标 t 代表年份。行业总需求因子的计算公式是: $q_{st} = \sum_{i=1}^{Ns} ms_{ist} \ln y_{ist}$, Ns 表示行业内企业的个数, y_{ist} 表示经价格指数平减后的企业销售收入。 $ms_{ist} = y_{ist} / \sum_{i=1}^{Ns} y_{ist}$ 表示行业内企业销售收入的份额。 $H_{it} = \delta_1 sc_{it} + \delta_2 export_{it} + \delta_3 young_{it} + \delta_4 soe_{it} + \delta_5 east_{it}$,为企业层面的一系列特征变量,控制了企业的横向产品差异。其中, sc_{it} 为企业营业费用占销售收入的比

① 不可观测或量化投入要素常见的例子是企业家精神、组织制度等。

② 以往的生产函数与成本函数一般的假设是产品市场是完全竞争的,企业在产品价格上没有控制力,然而现实中制造业往往呈现出产品差异化与多样化的特征。

重, $export_u$ 表示企业出口交货值占销售收入的比重, $young_u$ 为企业年龄是否小于 2 年的虚拟变量, soe_u 表示企业是否为国有企业的虚拟变量, $east_u$ 表示企业是否为东部地区的虚拟变量, δ 为待估参数。

在估计模型构建中,先设定总产值类型的成本函数,如式(16)。 C_u 表示企业产出变量为总产值时对应的总成本, y_u 表示总产值, P_{lu} 为劳动要素价格, P_{ku} 为资本要素价格, P_{mu} 为中间投入价格, β 为待估参数, ω_u 为成本函数下的 TFP, η_u 为测量误差。待估参数 $\beta_s = \frac{1}{|\eta_s|}$ 是对行业层面需求弹性的直接估计,控制了企业纵向的需求异质性, H_u 则是控制了企业的横向产品差异。

$$\ln C_u = \beta_y \ln y_u + \beta_l \ln P_{lu} + \beta_k \ln P_{ku} + \beta_m \ln P_{mu} + \beta_s q_{st} + H_u - \omega_u + \eta_u \quad (16)$$

假设企业 TFP 服从一阶马尔科夫过程 $\omega_u = g_t(\omega_{i,t-1}) + \xi_u$, ξ_u 表示 TFP 的外生波动。投资 I_u 与企业自身可观测到的 ω_u 具有严格单调递增的关系,并满足函数 $I_u = I(\omega_u, P_{ku}, P_{mu})$ 。对此求反函数可得:

$$\omega_u = I^{-1}(I_u, P_{ku}, P_{mu}) = \omega(I_u, P_{ku}, P_{mu}) \quad (17)$$

式(17)代入成本函数中可以得到半参数估计方程:

$$\begin{aligned} \ln C_u &= \beta_y \ln y_u + \beta_l \ln P_{lu} + \beta_s q_{st} + H_u + \beta_k \ln P_{ku} + \beta_m \ln P_{mu} - \omega(I_u, P_{ku}, P_{mu}) + \eta_u \\ &= \beta_y \ln y_u + \beta_l \ln P_{lu} + \beta_s q_{st} + H_u + \phi(I_u, P_{ku}, P_{mu}) + \eta_u \end{aligned} \quad (18)$$

$\phi(\cdot)$ 可以一个包含 $\ln I_u$ 、 $\ln P_{ku}$ 、 $\ln P_{mu}$ 的二阶多项式来近似。对式(18)使用最小二乘法(OLS)估计。半参数多项式控制了由于 TFP 波动与要素投入决策相关而导致的系数估计偏差。因此,误差项 η_u 不再与要素投入价格相关,控制了同步性问题, $\hat{\beta}_y$ 与 $\hat{\beta}_l$ 的系数估计具有无偏性和一致性。

再根据企业的退出决策行为,构造生存概率方程,以控制样本选择偏差问题。设置被解释变量为企业当期的生存决策 χ_u , 取决于企业的 TFP 阈值 ω_u 以及企业的经营状况信息 $J_{i,t}$ 。假设企业的 TFP 高于生存所需要最低程度的 TFP 时,企业的决策是继续生存下去,主要受上期的投资水平、资本价格、中间投入价格等因素影响。使用 Probit 模型来拟合企业生存概率 Pr,解释变量为 $\ln I_{i,t-1}$ 、 $\ln P_{ki,t-1}$ 、 $\ln P_{mi,t-1}$ 以及其平方项和交互项。拟合出来的企业生存概率为变量 \hat{P}_u 。

$$\begin{aligned} \Pr\{\chi_u = 1 \mid \omega_u(P_{ku}, P_{mu}), J_{i,t-1}\} &= \Pr\{\omega_u \geq \omega_u(P_{ku}, P_{mu}) \mid \omega_u(P_{ku}, P_{mu}), \omega_{i,t-1}\} \\ &= \rho_t\{\omega_u(P_{ku}, P_{mu}), \omega_{i,t-1}\} = \rho_t\{I_{i,t-1}, P_{ki,t-1}, P_{mi,t-1}\} \end{aligned} \quad (19)$$

将第一步得到的 $\hat{\beta}_y$ 与 $\hat{\beta}_l$ 回代到成本函数当中,构造非线性估计方程:

$$\begin{aligned} \ln C_u - \hat{\beta}_y \ln y_u - \hat{\beta}_l \ln P_{lu} - \hat{\beta}_s q_{st} - \hat{H}_u \\ = \beta_k \ln P_{ku} + \beta_m \ln P_{mu} - g(\hat{\phi}_{t-1} - \beta_k \ln P_{ki,t-1} - \beta_m \ln P_{mi,t-1}, \hat{P}_u) - \xi_u + \eta_u \end{aligned} \quad (20)$$

对式(19)使用非线性最小二乘法(NLS)估计, $g(\cdot)$ 使用包括 $\hat{\phi}_{t-1} - \beta_k \ln P_{ki,t-1} - \beta_m \ln P_{mi,t-1}$ 、 \hat{P}_u 的二阶多项式展开。该处理方式下,对资本价格弹性 β_k 的估计是无偏和一致的。在得到产出弹性与要素投入价格弹性的估计系数后,使用索洛余值法估算 TFP。构造 TFP 对数值的估计式:^①

$$\hat{\omega}_u = (\hat{\beta}_y \ln y_u + \hat{\beta}_l \ln P_{lu} + \hat{\beta}_k \ln P_{ku} + \hat{\beta}_m \ln P_{mu} + \hat{\beta}_s q_{st} + \hat{H}_u - \ln C_u) \left(\frac{\hat{\eta}_s}{\hat{\eta}_s + 1} \right) \quad (21)$$

① 参照 De Loecker(2011)的处理方式,由于函数设定中考虑了行业层面的需求异质性因素的影响,TFP 的估算需要使用行业产量的需求弹性进行调整。

增加值成本函数的估计方法与上述方法类似。区别在于增加值成本函数不考虑中间投入的影响,总成本变为对应的 C_{it}^{va} ,解释变量中的总产值 y_{it} 替换为增加值 va_{it} ,投资决策方程也相应变为 $I_{it}=I(\omega_{it},P_{kit})$ 。具体回归步骤与上述方法基本一致。^①

2. 样本选择与处理

本文选取中国工业企业数据库 1999—2007 年、2011—2013 年作为研究样本。为控制样本无效观测量的影响,剔除主要变量缺失和小于零的样本,包括销售收入、总产值、流动资产、固定资产、职工人数、中间投入等指标。剔除应交所得税缺失以及总资产低于固定资产等不符合会计基本准则的样本。为得到质量更高的企业数据,保留年销售收入 500 万元以上、职工人数不低于 8 人的样本,得到规模以上的工业企业数据。部分年份增加值数据缺失,参照聂辉华等(2012)的处理方法,1999—2003 年使用会计核算准则“工业增加值=工业总产值-工业中间投入+增值税”对增加值数据进行补充。2004 年工业总产值指标缺失,参照刘小玄和李双杰(2008)的估算方法“工业增加值=销售收入-期初存货+期末存货-工业中间投入+增值税”,对 2004 年增加值数据进行补足。最终得到 1999—2007 年间 556039 个工业企业观测量。

现有研究使用的 2009—2013 年的工业企业数据库样本,与 1999—2007 年部分相比存在较多的统计差异,包括关键变量诸如中间投入、增加值指标缺失。^② 部分年份亦出现其他核心指标的遗漏情况,如 2009—2010 年缺失应付工资总额数据,2010 年缺失折旧数据。已有相关研究均指出 2009 年、2010 年的统计数据质量较差,出现较为严重的法人代码缺失现象,样本统计质量难以保证(谭语嫣等,2017;余森杰等,2018)。综合考虑,采用 2011—2013 年数据作为研究样本。现对该部分样本中存在统计缺口的指标进行估算补足。使用样本中提供的应付工资总额与在职职工人数之比来表示人均工资。

中间投入的估算是重点也是难点,其不仅是总产值生产函数的重要投入要素,而且也是估算增加值的必要指标之一。由于 2011—2013 年关键变量(中间投入、增加值)缺失,需要重新估算,并以 1999—2007 年存在中间投入与增加值核算指标的数据作为基准评估估算质量。根据《工业企业统计报表制度》,工业中间投入包含五个部分,分别是直接材料、制造费用中的中间投入部分、营业费用中的中间投入部分、管理费用中的中间投入部分以及财务费用。工业中间投入通常有正算法和倒算法两种计算方法。由于 2008 年以后数据库变量缺失的问题,正算法与倒算法均缺乏必要的指标。^③ 在数据可得性的前提下,使用其他替代方法来估算工业中间投入也成为了更为合理的做法。余森杰等(2018)使用的估算公式是,“中间投入=总产值×主营业务成本/主营业务收入-应付工资总额-当年折旧”。该方法的优点在于对原始数据的要求较低,对主营业务成本部分通过使用总产值与主营业务收入调整为当期的成本支出,再剔除其中的属于劳动与资本要素的报酬,最终得到营业成本中

① 增加值成本函数的估计方法请参见《中国工业经济》网站(<http://www.ciejournal.org>)附件。

② 企业进入调查样本的门槛也发生了变化,2011 年对“规模以上”的统计口径进行调整,将进入统计样本的门槛从销售收入 500 万元上调至 2000 万元。

③ 正算法使用制造费用、营业费用、管理费用中属于中间投入的部分以及材料支出、利息支出的总和来测算工业中间投入。然而,2011—2013 年样本中,制造费用、营业费用、管理费用中属于中间投入的部分数据以及制造费用本身并未统计。这使得直接使用正算法来估算中间投入并不具备可行性。倒算法中,2011—2013 年样本仅报告了营业费用、管理费用,两类费用中福利费、保险费、办公费等非中间投入部分的支出指标缺失,导致使用倒算法估计得到的中间投入指标严重高估。本文在数据可得性的基础上使用倒算法估算 1999—2007 年的工业中间投入,发现与报告值在各类统计特征上相距甚远。

属于中间投入的部分。不足之处在于忽略了销售费用、管理费用中属于中间投入的部分以及忽略了将会全部计入中间投入的财务费用。对于中国工业企业的生产技术特征而言,中间投入部分主要体现在直接费用部分。因此,该方法具有一定的合理性。

本文将企业的财务费用考虑进来,在余淼杰等(2018)处理方法的基础上进行调整,构建估算公式“中间投入=总产值 \times 主营业务成本/主营业务收入-应付工资总额-当年折旧+财务费用”。以1999—2007年的工业企业数据为样本估算得到的中间投入均值是6568.370万元,同时期的中间投入原始数据均值为6519.746万元,二者基本接近,说明该方法有一定合理性。再根据会计核算准则“工业增加值=工业总产值-工业中间投入+增值税”对2011—2013年的工业增加值指标进行估算。

为了保证各类估计结果的可比性,剔除了估算后中间投入、增加值依旧缺失的样本,剔除关键变量小于0或缺失的样本,保留销售收入不低于2000万元,在职员工不低于8人的样本,筛选后最终得到2011—2013年有效的观测量628723个。变量构造上,依据生产函数与成本函数计量方程的标准,构造总产值和增加值与对应的总成本,并分别构造对应的要素投入量与对应的要素价格。^①

五、模拟分析

1. 拟蒙特卡洛实验设置

本文以中国工业企业数据库1999—2007年部分为模拟样本构造拟蒙特卡洛实验。剔除了“白噪声”的工业企业数据总体质量较好、可信度较高,相比随机生成数据的方法,能够一定程度上避免数据生成的随意性。由于现实中企业的生产行为十分复杂,使用现实的数据模拟能够较好地反映实际情形。通过每次随机抽取30%的样本,共抽样1000次,得到生产函数与成本函数的全部回归样本。在样本变量的选取上,抽取了数据库中的销售收入 y_{it} ,增加值 va_{it} ,资本投入 k_{it} ,劳动投入 l_{it} ,中间投入 m_{it} ,总产值成本 C_{it} ,增加值成本 C_{it}^{va} ,资本价格 P_{kit} ,劳动价格 P_{lit} ,中间投入价格 P_{mit} ,投资 I_{it} ,行业总需求因子 q_{st} ,横向产品差异因子 H_{it} 以及退出行为 $exit_{it}$ 的观测值作为拟蒙特卡洛实验的模拟样本。通过对每次抽样进行回归并提取回归系数得到生产函数与成本函数的系数估计结果,并运用索洛余值法计算TFP,提取TFP的均值与方差作为评价TFP测算结果的主要评价指标,TFP方差越大,意味着测算得到的TFP分布更为离散,异质性更强。鉴于OP估计方法的一个关键假设是TFP波动服从一阶马尔科夫过程,对 $\omega_{it}=\varphi\omega_{it-1}+v_{it}$ 式进行估计($|\varphi|<1$),测算得到估计系数 φ 。如果 φ 系数越大,则表明当期TFP受上一期TFP的影响越大,企业能够依据自身观察到的TFP变化来进行生产决策,其投资行为与TFP波动的单调关系更强。

2. 模拟结果分析

基于数据质量较好的工业企业样本(1999—2007年),进行拟蒙特卡洛实验,验证工业增加值偏差以及遗漏不可测生产要素的估计偏差,模拟结果如表1所示。TFP异质性的测算结果表明,生产函数中,增加值技术得到的TFP方差(0.6396)要远远高于总产值技术得到TFP方差(0.1286),验证了假说2。增加值技术未考虑到中间投入因素的影响,使得测算得到TFP存在两个方面的偏差:①生产要素弹性系数中缺失中间投入因素而导致的偏差,尽管增加值作为被解释变量,但未考虑到中间投入对增加值的影响,也会出现变量遗漏的问题。②未考虑到中间投入占产出的份额而导致的偏差,如果中间投入份额 S_{it} 是动态变化的,那么企业的增加值占总产值的比例($1-S_{it}$)也不是固定的

^① 变量构建方法与描述性统计请参见《中国工业经济》网站(<http://www.ciejournal.org>)附件。

常数。^① 缺失了中间投入变量的增加值生产函数, 计量方程中要素份额的动态变化会被归入到残差项中, 进而扩大了估计系数与 TFP 的估计偏差。而总产值生产函数将中间投入要素作为主要解释变量, 考虑并控制了中间投入占总产值比例变化的影响, 使其不会被归入到残差项中, 则回归结果更为稳健。对比要素投入的弹性系数也可以发现, 增加值技术下的资本弹性系数 $\ln k$ 与劳动弹性系数 $\ln l$ 的标准差均高于总产值生产函数, 表明增加值生产函数系数估计结果的异质性更强, 假说 1 得到支持。

表 1 拟蒙特卡洛实验模拟结果

增加值生产函数	均值	标准差	总产值生产函数	均值	标准差
$\ln k$	0.2620	0.0011	$\ln k$	0.0658	0.0005
			$\ln m$	0.7014	0.0016
$\ln l$	0.4044	0.0001	$\ln l$	0.0522	0.0000
TFP 均值	2.9007	0.0082	TFP 均值	2.1967	0.0151
TFP 方差	0.6396	0.0002	TFP 方差	0.1286	0.0005
φ	0.7453	0.0003	φ	0.6624	0.0022
模拟次数	1000		模拟次数	1000	
增加值成本函数	均值	标准差	总产值成本函数	均值	标准差
$\ln v a$	0.5512	0.0001	$\ln y$	0.7990	0.0000
$\ln P_k$	0.1536	0.0008	$\ln P_k$	0.1182	0.0004
			$\ln P_m$	0.3055	0.0003
$\ln P_l$	0.2777	0.0001	$\ln P_l$	0.1883	0.0001
TFP 均值	-3.1933	0.0025	TFP 均值	-1.6024	0.0014
TFP 方差	0.4349	0.0001	TFP 方差	0.1942	0.0001
φ	0.8135	0.0002	φ	0.8207	0.0002
模拟次数	1000		模拟次数	1000	

成本函数的估计结果表明, 增加值技术下 TFP 方差(0.4349)明显高于总产值技术下的 TFP 方差(0.1942), 假说 4 成立。类似地, 增加值成本函数的技术设定同样未考虑到中间投入因素, 使得测算得到 TFP 也存在两方面的偏差: ①成本函数要素价格的弹性系数中缺失中间投入价格而导致的偏差, 增加值成本函数未考虑中间投入价格变量导致遗漏变量的问题。②未考虑到中间投入占总产值以及总成本份额变化而导致的偏差。由于生产规模变化等原因, 各个生产要素份额会出现不同程度的改变, 中间投入占总成本的份额 S_{ci} 与占总产出的份额 S_{ii} 也会出现不同程度的改变。在产出弹性系数外生给定的条件下, 中间投入占总产出和总成本的份额的动态变化会导致增加值成本函数估计产生偏差。因此, 原本是要素份额的波动, 在增加值成本函数方程下会被错误地归入到残差项中, 进而加大了 TFP 估计的异质性。同样地, 增加值偏误造成的遗漏变量问题与中间投入份额的变化扩大了估计系数的波动性。对比两类成本函数的产出弹性系数与要素价格弹性系数可以发现, 增加值成本函数估计系数的标准差均高于总产值成本函数, 假说 3 得到支持。

3. TFP 均值分布分析

对 1000 次模拟得到的 TFP 均值分布进行偏度与峰度检验(Skewness and Kurtosis Test)以及单一样本 K-S 检验(Kolmogorov-Smirnov Test), 如表 2 所示。检验的原假设是“变量分布服从正态

^① 比如企业改变生产规模、改变多个最终产品的产出份额等因素会导致中间投入份额的变化。

分布”,若显著拒绝原假设,表明变量分布不服从正态分布,反之,则不能拒绝变量服从正态分布的原假设。根据各类 TFP 分布检验的 P 值,增加值生产函数与总产值生产函数均在 1%水平上显著拒绝原假设,其测算得到的 TFP 均值不服从正态分布。增加值成本函数在偏度与峰度检验中拒绝原假设,在 K-S 检验中不能拒绝变量服从正态分布的原假设。总产值成本函数在两类检验中均不能拒绝 TFP 均值服从正态分布的原假设,意味着成本函数测算得到的 TFP 要更接近于正态分布,测算结果更为稳健。

表 2 各类 TFP 均值的分布检验

变量	函数设定	观测量	偏度与峰度检验			K-S 检验	
			Pr(偏度)	Pr(峰度)	P 值	D 值	P 值
TFP 均值	增加值生产函数	1000	0.0000	0.0000	0.0000	0.1200	0.0000
TFP 均值	总产值生产函数	1000	0.0000	0.0000	0.0000	0.1337	0.0000
TFP 均值	增加值成本函数	1000	0.0034	0.1431	0.0061	0.0245	0.5860
TFP 均值	总产值成本函数	1000	0.7906	0.3228	0.5912	0.0198	0.8280

进一步对每次随机试验得到的 TFP 均值进行核密度估计,直观地比较 TFP 分布情况,如图 1 所示。相对于生产函数,成本函数估计得到的结果更为接近正态分布,表明成本函数估计结果更为稳定。生产函数无论是增加值形式还是总产值形式,都出现 TFP 分布的长尾效应,左端分布的低效率情况占比偏高,且分布密度也相应低于成本函数,导致 TFP 分布更为离散。其中的原因可能在于成本函数解释变量中自身含有产出项,能够更为有效地控制不可观测的其他生产要素(比如企业家精神、组织制度)的影响,验证了假说 5。生产函数存在着遗漏变量的问题,原本应该是由生产要素决定的产出部分,会被错误地归入到生产函数计量方程的残差项中,即 TFP 包含了其他生产要素的影响,这将导致测算得到的 TFP 不能够准确地反映企业当前的生产效率状态。同时,不可观测的生产要素的冲击使得 TFP 分布波动性更大,分布越发分散,与正态分布的特征大相径庭。

拟蒙特卡洛实验样本的观测期时间跨度较大,为 1999—2007 年,企业生产规模、投入结构甚至是产品价格水平均发生了较大变化,且其他未考虑到的生产要素投入亦发生了显著的改变。因此,在未能定量分析或者不可观测的要素投入未能考虑到时,生产函数的估计会出现一定的估计偏差,使得 TFP 模拟得到的分布呈现“长尾效应”。尽管成本函数计量方程的残差项也会包含不可观测的要素价格因素,但其自身拥有产出项作为解释变量,一定程度上稀释了不可观测生产要素对总成本的影响,能够较好地控制变量遗漏的问题,模拟得到的 TFP 更为稳定,分布更为接近正态分布。

正如理论分析所示,遗漏变量问题会导致方程估计产生显著的偏误,必然导致估计系数的不稳健。进一步分析导致生产函数分布不稳定的深层次原因,以生产函数与成本函数资本要素的弹性系数为例,对其进行核密度估计,刻画其估计系数的分布图。^①除增加值生产函数资本弹性系数以外,其余三类回归得到的估计系数核密度分布均相对集中,更为接近正态分布的形态。增加值偏差加大了计量方程系数的异质性,且生产函数的遗漏变量问题相对严重,导致增加值生产函数中资本投入的弹性系数呈现长尾效应,估计偏差更大。总之,在使用生产函数或成本函数等参数法来测算 TFP 的过程中,需要更为警惕由于遗漏重要生产要素变量带来的偏误,其根本原因在于遗漏的生产要素会被错误地归入到残差项中,使得原本相对“纯净”的 TFP 受到其他因素波动的影响,最终导致了研究结果的偏误。

^① 资本、中间投入与劳动要素估计系数的分布图请参见《中国工业经济》网站(<http://www.ciejournal.org>)附件。

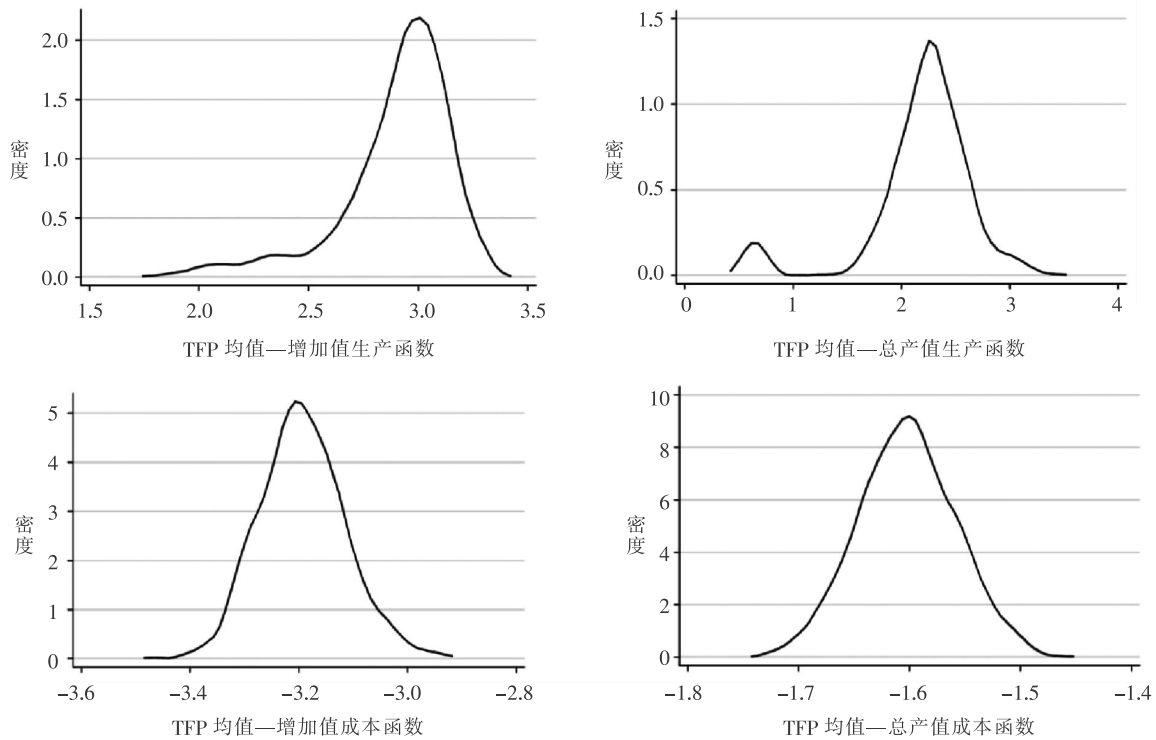


图1 TFP均值的核密度分布

六、实证分析

1. 基准回归结果

上文基于拟蒙特卡洛模拟方法对不同函数估算得到的TFP进行比对分析。为更为细化分析TFP所有制结构的分布差异以及动态演变的趋势,在全样本的视角下对1999—2007年的样本以及2011—2013年的样本分别进行回归,对四类TFP的测算结果进行分析。根据筛选得到的工业企业数据进行回归,结果如表3所示。1999—2007年样本中各个参数的估计结果与上文拟蒙特卡洛实验结果基本一致,表明拟蒙特卡洛实验的结果具有稳健性。选用筛选后的工业企业数据能够较好地控制其他外生变量的干扰,适用于分析特定技术设定下导致的TFP估算差异与异质性。

以1999—2007年样本的回归结果为基准,评估2011—2013年的回归质量。2011—2013年总产值生产函数与总产值成本函数的回归结果中,各类估计系数均接近于1999—2007年的回归结果,系数变化均在合理区间内,表明本文使用的中间投入与增加值估算方法具有合理性。反观增加值生产函数与增加值成本函数的估计结果,各类参数出现了较大变化,甚至出现较为不合理的结果。2011—2013年增加值生产函数中资本要素的估计系数相比1999—2007年出现了较大幅度的下降,增加值成本函数中资本投入价格也出现了低估的现象。造成上述估计偏差的原因,一方面是由于增加值的估算不能反映现实;另一方面增加值偏差的存在加大了增加值技术下参数估计的估计偏差,使得增加值函数的参数估计脱离了真实值。

2. TFP分布的所有制差异

前文对1999—2007年与2011—2013年的时间区间分别进行回归,在估算得到对应的TFP以后对企业所有制差异进行分析。一个值得探讨的问题是,国有企业领域不断深化改革的进程下,国

表 3 样本回归结果

增加值生产函数	估计值	估计值	总产值生产函数	估计值	估计值
$\ln k$	0.2513 (0.0050)	0.1800 (0.0055)	$\ln k$	0.0676 (0.0015)	0.0643 (0.0020)
			$\ln m$	0.6925 (0.0017)	0.7098 (0.0026)
$\ln l$	0.4044 (0.0023)	0.3186 (0.0023)	$\ln l$	0.0522 (0.0007)	0.0581 (0.0008)
时间	1999—2007	2011—2013	时间	1999—2007	2011—2013
增加值成本函数	估计值	估计值	总产值成本函数	估计值	估计值
$\ln va$	0.5513 (0.0011)	0.7319 (0.0011)	$\ln y$	0.7990 (0.0008)	0.7826 (0.0010)
$\ln P_k$	0.1661 (0.0039)	0.0282 (0.0032)	$\ln P_k$	0.1192 (0.0027)	0.1296 (0.0037)
			$\ln P_m$	0.3066 (0.0013)	0.2854 (0.0020)
$\ln P_l$	0.2777 (0.0022)	0.1360 (0.0011)	$\ln P_l$	0.1884 (0.0014)	0.1738 (0.0010)
时间	1999—2007	2011—2013	时间	1999—2007	2011—2013

注:括号内为估计值对应的标准误。

有企业表现出的经济绩效与民营企业之间的差异是否有所缩小?为此,在全样本回归的视角下,将企业分为国有企业与民营企业,比较不同所有制的企业 TFP 分布情况。^①

在具体的所有制定义上,排除了股权结构不明晰的企业、集体企业以及外资企业,仅考虑国有企业与民营企业的分布差异。根据资本占比确认企业性质,当国有资本占比超过 50%,则认定为国有企业;若私有资本占比超过 50%,则认定为民营企业。当国有资本与私有资本占比均为 50%时,则根据企业控股情况来判断企业的性质,若企业是国有控股时,则归类为国有企业,否则归类为民营企业。为了更好地对比国有企业与民营企业在不同时间段的效率差异,选用 2005—2007 年为第一段,2011—2013 年为第二段,分别测算二者的 TFP 分布,核密度估计结果如图 2 所示。

2005—2007 年,国有企业与民营企业在 TFP 分布上有着本质的区别,呈现出明显的所有制差异,整体上国有企业的 TFP 分布相比民营企业更为离散。2011—2013 年的结果则出现明显的不同,四类 TFP 的分布均表明,国有企业的 TFP 分布相当接近于民营企业。在混合所有制改革不断推进的背景下,国有企业的效率与民营企业相比差距明显缩小,国有企业深化改革增强了公有制经济的活力,市场在企业资源配置中的作用愈发重要。上述结果意味着,后金融危机时代以来,TFP 分布在所有制的差异方面发生了本质变化,国有企业与民营企业的 TFP 分布差异显著缩小,学术界在研究与评估混合所有制改革的政策效果时应予以重视。

值得注意的是,2005—2007 年各类 TFP 分布呈现出一定的差异。由于成本函数的被解释变量考虑了企业的代理成本,因此无论是总产值成本函数还是增加值成本函数,民营企业 TFP 分布的均值均要相对高于国有企业。国有企业高昂的代理成本导致了在经济运行效率上的损失。以样本中的管理费用为例,2005—2007 年民营企业的管理费用均值为 225 万元,国有企业管理费用支出均值

^① 为排除极端值对 TFP 分布的影响,对样本进行上下 0.5%的缩尾处理,下同。

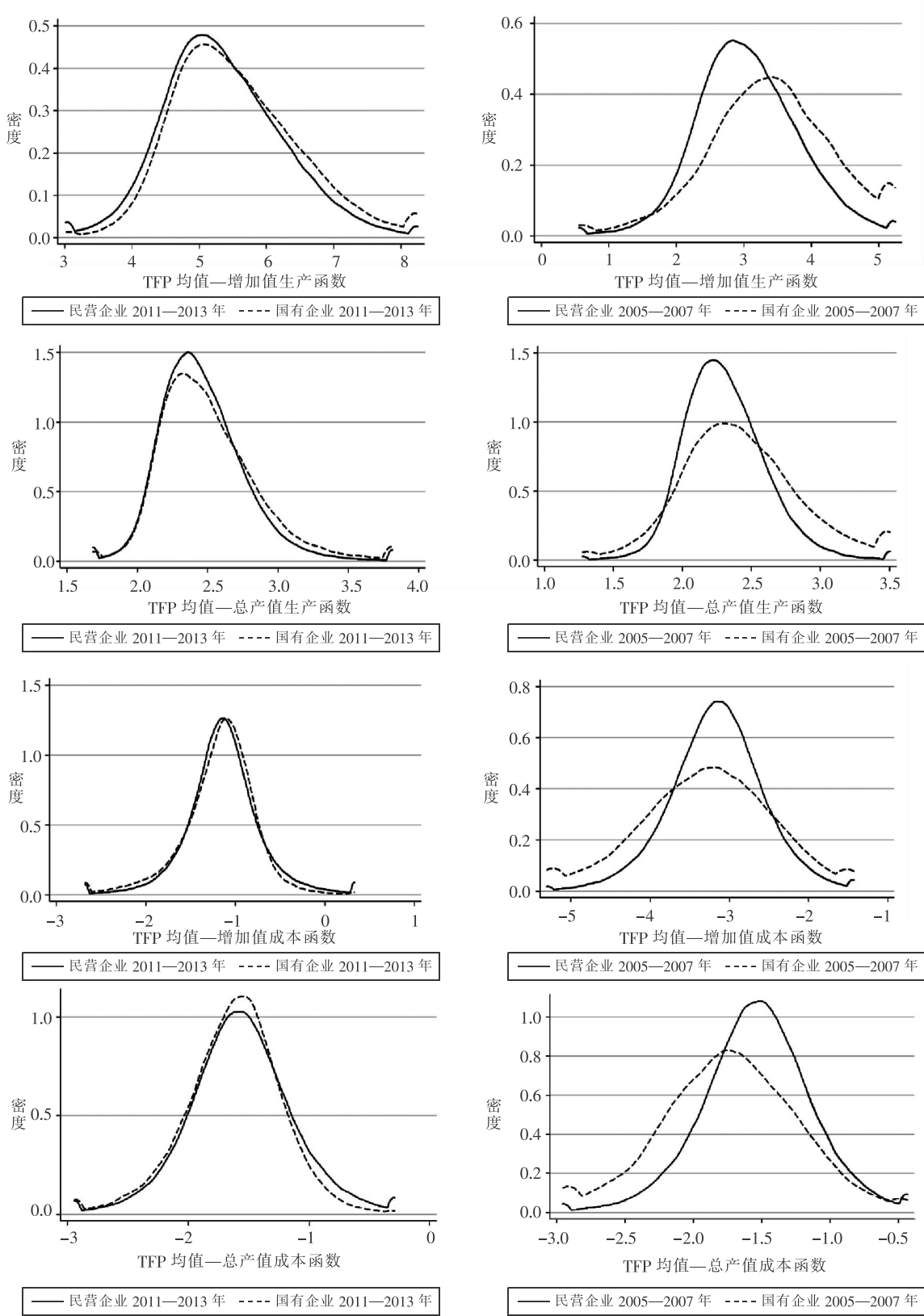


图2 各类TFP在不同时间区间分布的所有制差异

注：横轴是企业TFP，纵轴是分布密度。

高达 3462 万元。从份额看,民营企业样本中,管理费用占当年销售收入的份额约为 5%,国有企业管理费用的份额则高达 14%。根据成本函数的 TFP 计算公式,在既定的产出以及要素价格贡献下,国有企业偏高的管理费用支出将会加大企业成本,进而导致国有企业以成本视角核算的生产效率显著低于民营企业。^①相比之下,生产函数未考虑到代理成本的问题,产出项为企业实现的产出,要素投入项也未包括企业代理成本,因此,测算得到国有企业与民营企业的生产效率实际上并未扣除企业代理成本的影响,生产函数测算得到的 TFP 可能存在一定程度的高估。2011—2013 年,国有企业平均管理费用支出为 1612 万元,民营企业平均管理费用支出为 1069 万元,两类企业管理费用占当年销售收入份额均在 5% 的水平左右。这意味着随着国有企业领域市场化改革、混合所有制改革的推进,国有企业的代理成本大大降低,也在一定程度上解释了为何 2011—2013 年企业 TFP 分布的所有制差异明显缩小。

3. 行业层面的配置效率变化:TFP 动态分解

进一步对行业层面的 TFP 演进趋势进行动态分解,以比较不同方法测算的 TFP 在研究资源配置时导致的结论差异。根据 Melitz and Polanec(2015)提出的动态分解方法(简称 MP 方法),TFP 的变动 $\Delta\Phi$ 可以分解为四个部分,分别是生存企业的技术进步 $\Delta\bar{\omega}_s$ 、生存企业间的资源配置效率 Δcov_s 、高效率企业进入 $S_{E2}(\Phi_{E2}-\Phi_{S2})$ 与低效率企业退出 $S_{X1}(\Phi_{S1}-\Phi_{X1})$,如式(22)所示:^②

$$\Delta\Phi = \Delta\bar{\omega}_s + \Delta cov_s + S_{E2}(\Phi_{E2} - \Phi_{S2}) + S_{X1}(\Phi_{S1} - \Phi_{X1}) \quad (22)$$

根据 MP 分解方法,测算了 1999—2007 年以及 2011—2013 年各类 TFP 的动态分解结果,如表 4 所示。1999—2007 年,增加值生产函数测算得到 TFP 年均变动 0.0814,明显高出总产值生产函数以及两类成本函数的估算结果,表明增加值生产函数方法对应的 TFP 变动被明显地高估,也侧面印证了增加值生产函数的设定导致了 TFP 估计偏差的理论假说。由于中间投入要素的贡献纳入到作为残差项的 TFP 中,将会增大 TFP 的异质性,使得 TFP 分布更为离散,因此,增加值生产函数的 TFP 进行动态分解后,各部分贡献会出现较大偏差的结果。类似地,增加值成本函数同样存在着遗漏中间投入要素的问题,使其估算得到的 TFP 总变动尽管与总产值成本函数的结果相差不大,然而在全类细分部分的估算结果均存在着一定程度的偏离,导致在评估企业间、进入企业等部分的贡献率时得到不稳健的结论。换言之,增加值偏差的存在使增加值生产函数分解得到的 TFP 增长速度被明显高估,且 TFP 分解后各部分贡献出现较大的测算误差,导致不稳健的测算结果。研究者在进行 TFP 动态演变的相关分析时,需要注意由于增加值偏差导致的 TFP 异质性问题。

2011—2013 年,除了增加值生产函数得到的 TFP 总变动出现较大幅度的效率下滑以外,其他三类 TFP 的变动均呈现停滞甚至是一定程度下滑的趋势,与 2008 年以前的 TFP 普遍呈现增长趋势的结论截然不同。后金融危机时代以来,尽管中国经济仍然呈现高速增长态势,但制造业企业的 TFP 并未得到相应改善,针对工业企业效率下降的原因与对策的研究则具有重要的现实意义与政策启示。

从行业层面 TFP 演变的视角出发,选取纺织(17)、食品制造(14)、木材加工(20)、造纸(22)、石油及核燃料加工(25)、化工原料制品(26)、非金属矿物制品(31)、通用设备制造(35)共 8 个两位数代码行业测算配置效率的变化。以增加值生产函数与总产值生产函数估算得到的 TFP 为例,MP 方

^① 现有文献通常采用工业企业的管理费用来代理企业的代理成本,并用于研究国有企业改革、行政垄断的绩效评估中,详见孔东民等(2014)研究结果。

^② 具体方法请参见《中国工业经济》网站(<http://www.ciejjournal.org>)附件。

表 4 TFP 分解结果

时间	总体变化 $\Delta\Phi$	企业内 $\Delta\bar{\omega}_s$	企业间 Δcov_s	生存企业 $\Delta\bar{\omega}_s + \Delta cov_s$	退出企业 $S_{E2}(\Phi_{E2} - \Phi_{S2})$	进入企业 $S_{X1}(\Phi_{S1} - \Phi_{X1})$
测度方法:增加值生产函数						
1999—2007	0.0814	0.0530	0.0337	0.0867	0.0048	-0.0102
2011—2013	-0.0397	0.0456	-0.0667	-0.0211	0.1619	-0.1805
测度方法:总产值生产函数						
1999—2007	0.0303	0.0254	0.0121	0.0375	0.0018	-0.0090
2011—2013	-0.0206	0.0182	-0.0348	-0.0166	0.0707	-0.0747
测度方法:增加值成本函数						
1999—2007	0.0222	0.0287	-0.0253	0.0033	-0.0059	0.0248
2011—2013	0.0019	-0.0496	0.0533	0.0037	-0.0493	0.0475
测度方法:总产值成本函数						
1999—2007	0.0351	0.0276	-0.0055	0.0221	-0.0037	0.0167
2011—2013	0.0052	-0.0559	0.0643	0.0084	-0.0553	0.0521

注:在此展示的是 1999—2007 年和 2011—2013 年 TFP 的年平均变化均值。

法分解结果如表 5 所示。^① 可知,尽管在 1999—2007 年,各个子行业(除石油及核燃料加工(25)外)的 TFP 呈现出改进的趋势。在次贷危机冲击后,随着全球增速放缓,各个行业在 2011—2013 年的 TFP 变化上出现不同程度上的分化。以总产值生产函数测算得到 TFP 分解结果为基准,2011—2013 年,食品制造、纺织、木材加工等行业均呈现出较大幅度的效率下滑,主要原因在于次贷危机影响下外部需求不足使得企业的生产效率逐步下滑。造纸、化工原料制品等制造业的配置效率由于产能过剩等因素呈现出停滞的趋势。通用设备制造、石油及核燃料加工、非金属矿物制品等为代表的制造业则随着高技术产业的发展呈现出 TFP 提升甚至是大程度改进的趋势。TFP 分解的结果表明,国际环境不确定性不断增大,中国政府需要继续深化推进市场化改革、改善企业间的资源配置效率,使要素资源能够有效地向高效率企业倾斜,切实降低企业间的资源错配现象,进一步深化经济体制领域的改革,以提高企业全要素生产率、优化经济结构、促进经济发展。

七、主要结论与启示

本文基于增加值偏差的视角,结合理论分析与实证分析,比较了由于产出变量设定差异而导致的估计偏差。以中国工业企业数据为样本构造拟蒙特卡洛实验,比较了在增加值生产函数、总产值生产函数、增加值成本函数与总产值成本函数下的 TFP 估计结果。研究发现:①增加值生产函数测算得到的 TFP 异质性要大于总产值生产函数,增加值成本函数测算得到的 TFP 异质性要大于总产值成本函数。原因在于,增加值产出的函数设置忽略了中间投入因素的影响。随着企业改变生产规模,中间投入占总产出与总成本的份额并非一成不变。如果不考虑中间投入的影响,那么这些偏差将会被识别到残差项里,最终加大了 TFP 的异质性。基于此,总产值类型的函数相比增加值函数的稳健性更好。②从遗漏难以观测变量的角度看,生产函数测算得到的 TFP 分布出现长尾效应,主要源于 TFP 以外其他不可观测的生产要素出现变化时会直接加大残差项的离散程度,TFP 的异质性更大。而由于成本函数的解释变量中含有产出项,有效稀释了由于遗漏变量的变化带来的偏差,成本函数测算得到的 TFP 相对生产函数更为接近正态分布。③以生产函数与成本函数中的资本要素

① 详细的分行业 TFP 动态分解结果请参见《中国工业经济》网站(<http://www.ciejournal.org>)附件。

表 5 代表性行业 TFP 分解结果

行业	时间	TFP 总体变化	
		增加值生产函数	总产值生产函数
食品制造(14)	1999—2007	0.0886	0.0305
	2011—2013	-0.0625	-0.0216
纺织(17)	1999—2007	0.0693	0.0319
	2011—2013	-0.0779	-0.0400
木材加工(20)	1999—2007	0.0944	0.0266
	2011—2013	-0.0479	-0.0284
造纸(22)	1999—2007	0.0733	0.0276
	2011—2013	0.0282	0.0006
石油及核燃料加工(25)	1999—2007	-0.0716	0.0264
	2011—2013	0.1611	0.0855
化工原料制品(26)	1999—2007	0.1031	0.0451
	2011—2013	-0.0472	-0.0071
非金属矿物制品(31)	1999—2007	0.0856	0.0347
	2011—2013	0.8318	0.4152
通用设备制造(35)	1999—2007	0.0925	0.0323
	2011—2013	0.0489	0.0198

注:在此展示的是 1999—2007 年和 2011—2013 年 TFP 的年变化均值。

的弹性系数模拟结果为例,在增加值偏差与遗漏不可观测变量问题的双重效应下,增加值生产函数得到的资本弹性系数分布更为离散,进而导致测算得到的 TFP 异质性更大。

以 1999—2013 年工业企业数据为样本,估算企业在 2008 年前后的 TFP 分布差异与动态演变趋势。研究发现:一方面,2011—2013 年企业 TFP 分布的所有制差异明显缩小,国有企业与民营企业的 TFP 分布差异比 2005—2007 年有着明显的减少,表明国有企业混合所有制改革、市场化改革有效减低了所有制差异,改善了国有企业的绩效;另一方面,行业层面的 TFP 动态分解结果显示,与 1999—2007 年行业 TFP 呈现提升的趋势相比,2011—2013 年行业 TFP 出现一定程度上的停滞甚至是下降的趋势。针对后金融危机时代的工业企业 TFP 出现的新特征,不仅需要考虑到 TFP 的估算问题,更要结合新的经济形势变化进行更为深入全面的研究。

本文的主要结论表明增加值与总产值技术的测算结果具有本质的区别,为微观企业的 TFP 研究提供有益的启示。

(1)全要素生产率是衡量微观企业绩效的核心指标,其本身不可观测,只能通过间接的方法来进行测算得到。因此,需要考虑到由于计量方法设置不当而可能导致的全要素生产率异质性问题,且必须考虑到使用增加值作为产出项引致的全要素生产率估计偏误。在数据可得的前提下,应尽可能使用多种方法对全要素生产率进行测算和结果对比,力求准确地评价企业的生产效率,避免可能出现的统计问题以及导致的结论偏误。

(2)相比以往的生产函数理论研究,成本函数方法论能够有效补充和完善全要素生产率的测算体系。打开企业生产的“黑箱”不仅需要借助生产函数工具,更需要成本函数作为辅助工具才能更全面地刻画企业生产规律。在后续的全要素生产率研究中,成本函数方法测算得到的企业绩效将具有重要的参考价值。

(3)随着全球经济下行承压,国际环境不确定性不断加剧,制造业的产业格局与金融危机前产生了本质的变化。及时测度与合理评价企业的全要素生产率,识别在金融危机以来中国制造业资源

配置、效率动态演变趋势,能为解释和探索制造业产业升级的脉络和规律提供有益的现实参考。相应地,中国制造业面临技术更新换代的转型压力与发达国家高端制造业的竞争压力。深化混合所有制改革、完善国有资产管理体制、引导和加速制造业企业的技术迭代,有利于中国在全球经济格局的演变中占领战略高地,进而提升中国制造业出口在全球价值链的位置。不可否认的是,使用传统的经济增速指标不再适用于评价中国高质量发展的绩效。因此,科学地评价工业经济高质量发展水平,重视微观企业全要素生产率测算所需要注意的问题,构建更为准确合理的效率评价指标,对提高供给侧结构性改革绩效与实现经济高质量发展具有重要的指导意义。

[参考文献]

- [1]陈林. 自然垄断与混合所有制改革——基于自然实验与成本函数的分析[J]. 经济研究, 2018, (1):81-96.
- [2]贺胜兵,周华蓉. 国际金融危机背景下国有和私营工业企业 TFP 增长的比较研究[J]. 系统工程, 2014, (8):18-25.
- [3]黄先海,金泽成,余林徽. 要素流动与全要素生产率增长:来自国有部门改革的经验证据[J]. 经济研究, 2017, (12):62-75.
- [4]孔东民,代昫昊,李阳. 政策冲击、市场环境与国企生产效率:现状、趋势与发展[J]. 管理世界, 2014, (8):4-17.
- [5]柳荻,尹恒. 企业全要素生产率估计新方法——全要素生产率估计的结构方法及其应用[J]. 经济学动态, 2015, (7):136-148.
- [6]刘小玄,李双杰. 制造业企业相对效率的度量和比较及其外生决定因素(2000—2004)[J]. 经济学(季刊), 2008, (3):843-868.
- [7]鲁晓东,连玉君. 中国工业企业全要素生产率估计:1999—2007[J]. 经济学(季刊), 2012, (2):541-558.
- [8]聂辉华,贾瑞雪. 中国制造业企业生产率与资源误置[J]. 世界经济, 2011, (7):27-42.
- [9]聂辉华,江艇,杨汝岱. 中国工业企业数据库的使用现状和潜在问题[J]. 世界经济, 2012, (5):142-158.
- [10]谭语嫣,谭之博,黄益平,胡永泰. 僵尸企业的投资挤出效应:基于中国工业企业的证据[J]. 经济研究, 2017, (5):175-188.
- [11]王永进,冯笑. 中国混合出口企业生产率研究:经验事实与理论解释[J]. 世界经济, 2019, (7):71-93.
- [12]王永进,刘灿雷. 国有企业上游垄断阻碍了中国的经济增长?——基于制造业数据的微观考察[J]. 管理世界, 2016, (6):10-21.
- [13]许明,李逸飞. 最低工资政策、成本不完全传递与多产品加成率调整[J]. 经济研究, 2020, (4):167-183.
- [14]杨汝岱. 中国制造业企业全要素生产率研究[J]. 经济研究, 2015, (2):61-74.
- [15]尹恒,柳荻,李世刚. 企业全要素生产率估计方法比较[J]. 世界经济文汇, 2015, (4):1-21.
- [16]尹恒,张子尧. 需求异质与企业加成率估计[J]. 中国工业经济, 2019, (12):60-77.
- [17]余森杰,金洋,张睿. 工业企业产能利用率衡量与生产率估算[J]. 经济研究, 2018, (5):56-71.
- [18]于永达,吕冰洋. 中国生产率争论:方法的局限性和结论的不确定性[J]. 清华大学学报(哲学社会科学版), 2010, (3):141-153.
- [19]张志强. 微观企业全要素生产率测度方法的比较与应用[J]. 数量经济技术经济研究, 2015, (12):107-123.
- [20]周黎安,张维迎,顾全林,汪淼军. 企业生产率的代际效应和年龄效应[J]. 经济学(季刊), 2007, (4):1297-1318.
- [21]Akerberg, D. A., K. Caves, and G. Frazer. Identification Properties of Recent Production Function Estimators[J]. *Econometrica*, 2015, 83(6):2411-2451.
- [22]Bloom, N., M. Schankerman, and J. Van Reenen. Identifying Technology Spillovers and Product Market Rivalry[J]. *Econometrica*, 2013, 81(4):1347-1393.
- [23]Brandt, L., C. T. Hsieh, and X. Zhu. China's Great Economic Transformation: Growth and Structural Transformation in China[M]. Cambridge:Cambridge Books, 2008.
- [24]Brandt, L., J. Van Biesebroeck, and Y. Zhang. Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing[J]. *Journal of Development Economics*, 2012, 97(2):339-351.

- [25]De Loecker, J. Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity[J]. *Econometrica*, 2011,79(5):1407–1451.
- [26]Doraszelski, U., and J. Jaumandreu. R&D and Productivity: Estimating Endogenous Productivity [J]. *Review of Economic Studies*, 2013,80(4):1338–1383.
- [27]Gandhi, A., S. Navarro, and D. A. Rivers. On the Identification of Production Functions: How Heterogeneous Is Productivity[R]. CIBC Working Paper Series, 2016.
- [28]Grieco, P. L. E., and R. C. McDevitt. Productivity and Quality in Health Care: Evidence from the Dialysis Industry[J]. *Review of Economic Studies*, 2016,84(3):1071–1105.
- [29]Levinsohn, J., and A. Petrin. Estimating Production Functions Using Inputs to Control for Unobservables[J]. *Review of Economic Studies*, 2003,70(2):317–341.
- [30]Melitz, M. J., and S. Polanec. Dynamic Olley–Pakes Productivity Decomposition with Entry and Exit [J]. *Rand Journal of Economics*, 2015,46(2):362–375.
- [31]Olley, S., and A. Pakes. The Dynamics of Productivity in the Telecommunications Equipment Industry[J]. *Econometrica*, 1996,64(6):1263–1297.
- [32]Syverson, C. What Determines Productivity[J]. *Journal of Economic Literature*, 2011,49(2):326–65.
- [33]Van Beveren, I. Total Factor Productivity Estimation: A Practical Review [J]. *Journal of Economic Surveys*, 2012,26(1):98–128.

Industrial Added Value and Total Factor Productivity Estimation——Evidence from the Quasi–Monte–Carlo Experiment of China’s Manufacturing Industry

ZHU Pei-hua¹, CHEN Lin²

- (1. School of Economics, Renmin University of China, Beijing 100872, China;
2. Institute of Industrial Economics, Jinan University, Guangzhou 510632, China)

Abstract: The Total Factor Productivity is one of the main statistical indicators of the modern economic system. However, there are many statistical errors needed to be corrected. And the estimation problem caused by using industrial value added as the proxy variable of output (relative to the gross output) is particularly serious. Therefore, this paper analyses the effect of the value-added bias on the results of productivity estimation by the mathematical model. In terms of empirical analysis, this paper establishes the gross-output and value-added production function, gross-output and value-added cost function model respectively, and estimates them by semi-parametric estimation method. The results of Quasi-Monte-Carlo experiment show that, due to the value-added bias, the productivity heterogeneity based on the value-added production function is greater than the gross-output production function, and the productivity heterogeneity calculated by the value-added cost function is larger than the gross-output cost function. The estimation results obtained by using gross output as the output proxy variable are more robust. According to the empirical results of China’s industrial enterprise data from 1999 to 2013, the ownership difference of TFP distribution from 2011 to 2013 was significantly reduced, compared with that before 2008. The results of TFP dynamic decomposition show that from 2011–2013, the productivity changes are stagnant or even declining, and vary among different industries. In order to get more realistic conclusions, it is imperative to pay attention to the new characteristics of productivity distribution and the new trends in productivity changes.

Key Words: total factor productivity; industrial value-added value; TFP estimation bias; value-added bias; cost function

JEL Classification: D24 L11 C52

[责任编辑:许明]